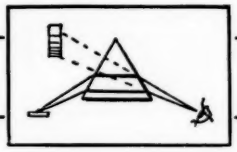
Journal of the OPTICAL SOCIETY *of* AMERICA

VOLUME 26
No. 11



NOVEMBER
1936

First-Order Laws in Asymmetrical Optical Systems. II. The Image Congruences Belonging to the Rays Emerging from a Point in Object and Image Space; Fundamental Forms . . . . . M. HERZBERGER 389

Effect of Pre-Exposure in Spectrum Photography . . . DANIEL NORMAN 407

Changes in Color Temperature of Tungsten-Filament Lamps at Constant Voltage . . . . . DEANE B. JUDD 409

Estimation of Chromaticity Differences and Nearest Color Temperature on the Standard 1931 ICI Colorimetric Coordinate System . DEANE B. JUDD 421

Published monthly
at Prince and Lemon Sts., Lancaster, Pa.
by the
AMERICAN INSTITUTE OF PHYSICS
Incorporated
175 Fifth Avenue, New York, N. Y.
FOR AND UNDER THE EDITORIAL SUPERVISION OF
THE OPTICAL SOCIETY OF AMERICA

THE OPTICAL SOCIETY OF AMERICA

OFFICERS OF THE SOCIETY

A. C. HARDY, *President*
Massachusetts Institute of Technology
Cambridge, Massachusetts

R. C. GIBBS, *Vice President*
Department of Physics
Cornell University
Ithaca, New York

L. B. TUCKERMAN, *Secretary*
Bureau of Standards
Washington, D. C.

HENRY F. KURTZ, *Treasurer*
Bausch & Lomb Optical Company
Rochester, New York

The Board of Directors consists of the above officers; the Editor of the Journal; the Past President; and five elected members.

W. B. RAYTON, *Past President*
Bausch & Lomb Optical Company
Rochester, New York

F. K. RICHTMYER, *Editor*
Cornell University
Ithaca, New York

Elected Members

ARTHUR H. COMPTON
Ryerson Physical Laboratory
University of Chicago
Chicago, Illinois

K. S. GIBSON
Bureau of Standards
Washington, D. C.

G. R. HARRISON
Department of Physics
Massachusetts Institute of Technology
Cambridge, Massachusetts

HERBERT E. IVES
463 West Street
New York, New York

CARL W. KEUFFEL
Keuffel & Esser Company
Hoboken, New Jersey

ROCHESTER, NEW YORK SECTION OF THE OPTICAL SOCIETY OF AMERICA

L. A. DUBRIDGE, *President*
F. M. E. HOLMES, *Secretary*

B. NOYES, *Vice President*
C. A. MORRISON, *Treasurer*

Councilors: F. E. ALTMAN, F. C. FAIRBANKS, L. V. FOSTER, O. SANDVIK, C. C. NITCHIE (*ex-officio*)

THE OPTICAL SOCIETY OF AMERICA was organized in 1916 "*to increase and diffuse the knowledge of optics in all its branches, pure and applied, to promote the mutual interests of investigators of optical problems, of designers, manufacturers and users of optical instruments and apparatus of all kinds and to encourage cooperation among them.*" It aims to serve the interests of all who are engaged in any branch of optics from fundamental research to the manufacture of optical goods.

MEMBERSHIP

Regular Members: "Any person who has contributed materially to the advancement of optics is eligible for regular membership."

Associate Members: "Any person who is interested in optics is eligible for associate membership."

Corporate Members: "Any corporation desirous of encouraging the work of the Society is eligible for corporation membership."

The Society cordially invites to membership all who are interested in any branch of optics, either in research, in instruction, in optical or illuminating engineering, in the manufacture and distribution of optical goods of all kinds or in physiological and medical optics.

ANNUAL DUES

Regular members, $7.50; Associate members, $5.00; Corporate members, $50.00. These dues include subscription to the Society's official journal, the *Journal of the Optical Society of America.*

MEETINGS

The Society holds two meetings per year: An annual meeting in the autumn and a midwinter meeting jointly with the American Physical Society.

For further information concerning the Optical Society of America address the Secretary, Dr. L. B. Tuckerman, National Bureau of Standards, Washington, D. C.

Journal of the
OPTICAL SOCIETY
of AMERICA

VOLUME 26 NOVEMBER, 1936 NUMBER 11

First-Order Laws in Asymmetrical Optical Systems

II. The Image Congruences Belonging to the Rays Emerging from a Point in Object and Image Space; Fundamental Forms*

M. HERZBERGER, *Eastman Kodak Laboratories, Rochester, New York*

(Received September 19, 1936)

I. INTRODUCTION

THE author thinks that the problems with which the present paper deals hold the key toward a better knowledge and a better understanding of the problems of optical image formation and especially of image errors. He will explain the reasons for his belief in the following. The theory of the first-order laws in asymmetrical systems has been given in its fundamentals by Allv. Gullstrand. T. Smith, in England, and H. Boegehold and the author, in Germany, have tried to simplify Gullstrand's results. The author has published in 1936 an explicit monograph[1] on the subject with a critical historical appendix.

There are three reasons why the author resumes work on the subject:

(1) The author has in the meantime discovered his tetrality principle,[2] with the help of which he was able to give an exhaustive survey of the problems of Gaussian optics. The methods and results given there apply in the same way to the more general problems of asymmetric optical systems.

The tetrality principle enables us to deal with laws of an infinite point as if it were a finite point, and gives the desirable generalization, whereas the methods previously used always allowed exceptions.

(2) The author uses in this presentation the Schleiermacher equations instead of the so-called Bruns equations. He hopes that the simplification of these equations will permit the development, later on, of an image error theory better suited for practical purposes. Let us sketch the difference between the Bruns and Schleiermacher equations. Hamilton discovered the eiconal function. Let us choose an arbitrary origin in object and image space and a Cartesian coordinate system, x, y, z and x', y', z', respectively, in each of them. Then, a ray is given by the coordinates x and y of the point in which it meets the plane $z=0$, and its direction cosines ξ, η (or better $n\xi$, $n\eta$; n being the refractive index in object space) with the x and y axes. The image ray is analogously given by x', y', $n'\xi'$, $n'\eta'$. Bruns found now, rediscovering and specializing the fundamental concept of Hamilton, the following theorem. If we choose from the eight coordinates, x, y, x', y', $n\xi$, $n\eta$, $n'\xi'$, $n'\eta'$, four, such that two belong to the object space and two to the image space, and assume that every ray which we consider is unambiguously defined by our choice of coordinates, then there exists al-

* Communication No. 593 from the Kodak Research Laboratories. For Part I, see J. O. S. A. **26**, 354 (1936).

[1] Zeits. f. Instrumentenk. **54**, 336–350, 381–392, 429–441 (1934).

[2] J. O. S. A. **25**, 295–304 (1935).

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

ways a function of these four coordinates such that the other four coordinates are the corresponding derivatives of this function (eventually with the negative sign). This function, called *eiconal* by Bruns, gives us all the qualities of the system, and from the existence of this function can be developed all optical laws. For instance, if we take $n\xi$, $n\eta$, $n'\xi'$, $n'\eta'$ as variables, we can find a function $W(n\xi, n\eta, n'\xi', n'\eta')$ such that the coordinates of the points where they meet the coordinate planes are given by

$$\begin{aligned} x &= \partial W/\partial n\xi, & x' &= \partial W/\partial n'\xi', \\ y &= \partial W/\partial n\eta, & y' &= \partial W/\partial n'\eta'. \end{aligned} \qquad \text{(I, 1)}$$

H. Boegehold and the author, in Germany, and T. Smith, in England, developed the laws of optics resulting from these and the corresponding formulae, and succeeded in finding for many special image formations, important from a theoretical or practical standpoint, the explicit form of the eiconal.[3] There is one difficulty, however, in applying these methods practically. What we want to know, and the only thing we can obtain from our numerical calculations, is the position of the image ray if the object ray is given, and *vice versa*. We want to find x', y', $n'\xi'$, $n'\eta'$ as functions of x, y, $n\xi$, $n\eta$; and x, y, $n\xi$, $n\eta$ if x', y', $n'\xi'$, $n'\eta'$ are given. That leads, if we use the Bruns equations, to a difficult problem of elimination which has not yet been successfully attacked. Another difficulty is the following. We know one of the eiconals for a single surface. We know that the eiconal for a composed system is equal to the sum of the eiconals of the composing systems but the actual calculation of the eiconal for a composed system would require the elimination of all the intermediate variables, since the eiconal contains only the coordinates in object and image spaces. It can be demonstrated that this elimination is possible, and the equations between which the eliminations must be made can be put into simple form; but the actual elimination, even with a kind of successive approximation, has not yet been achieved. Thus, the author tries to follow another line of investigation given by the Schleiermacher equations. The problem is to give x', y', $n'\xi'$, $n'\eta'$ as functions of x, y, $n\xi$, $n\eta$. x', y', $n'\xi'$, $n'\eta'$ must be unambiguous continuous functions of the object coordinates, which can be developed in Taylor series. These equations

$$\begin{aligned} x' &= f_1(x, y, n\xi, n\eta), \\ y' &= f_2(x, y, n\xi, n\eta), \\ n'\xi' &= f_3(x, y, n\xi, n\eta), \\ n'\eta' &= f_4(x, y, n\xi, n\eta), \end{aligned} \qquad \text{(I, 2)}$$

have been called Schleiermacher equations by the author because Schleiermacher gave for the first approximation and the special case of Gaussian optics, the necessary conditions between the coefficients. For, if we develop $f_1 \cdots f_4$ into a Taylor series, we find 16 first-order coefficients and 40 second-order coefficients, which are, of course, not all independent of each other.

In Part I of this paper we gave six conditions between the first-order coefficients, reducing them to ten independent coefficients. We shall restrict ourselves here to the consideration of these coefficients and their relations.

These Schleiermacher equations have two advantages over the Bruns equations: It is these equations which are obtained in calculating rays through an optical system, and they are, contrary to the Bruns equations, valid without exception for every optical system. A future development of optics certainly should make use of both types of equations: the elegant Bruns-Hamilton development to find the laws, and afterwards the Schleiermacher equations to find the connections with the actual calculations. This paper will give the first-order theory from the standpoint of the Schleiermacher equations.

(3) We will suggest in this paper standard forms of the Schleiermacher equations for each special case. These equations assume that the origins in object and image space lie on the principal ray, and that the z and z' axes coincide with it. They do not make any suggestion as to where to put the origin nor how to choose the coordinate planes. In this paper, we shall try to give in each case the simplest possible form to the Schleiermacher equations. This enables us, as in Gaussian optics, to develop all the laws for systems of special types, and the author believes that many of these laws will be considered interesting and useful.

Now, the question might be asked why this investigation has been made so general, whereas

[3] J. O. S. A. **26**, 5 (1936).

the chief interest of optical designers lies in rotation-symmetrical systems, where a certain kind of symmetry can be assumed. The answer is that even the laws of rotation-symmetrical systems cannot be understood without a thorough knowledge of the laws developed here. For instance, if we would study the second-order laws of these systems, the so-called image error theory, this theory must include the first-order laws for any neighboring ray of the optical axis, even for a skew ray. And the first-order laws for a skew ray are these very general laws of Gullstrand, which are here symmetrically developed and supplemented.

II. Variation of the Schleiermacher Equations with Changes of Origins and Coordinate Planes

A principal ray is given. We shall investigate the rays in its neighborhood, and in this second part, especially the neighboring rays which proceed from the points of the principal ray.

We choose the origins on the principal ray, the z, z' axes lying along it, and get

$$\begin{aligned} x' &= a_{11}x + a_{12}y + b_{11}n\xi + b_{12}n\eta, \\ y' &= a_{21}x + a_{22}y + b_{21}n\xi + b_{22}n\eta, \\ n'\xi' &= c_{11}x + c_{12}y + d_{11}n\xi + d_{12}n\eta, \\ n'\eta' &= c_{21}x + c_{22}y + d_{21}n\xi + d_{22}n\eta, \end{aligned} \tag{II, 1}$$

or, in matrix notation,

$$\begin{aligned} \begin{pmatrix} x' \\ y' \end{pmatrix} &= A\begin{pmatrix} x \\ y \end{pmatrix} + B\begin{pmatrix} n\xi \\ n\eta \end{pmatrix}, \\ \begin{pmatrix} n'\xi' \\ n'\eta' \end{pmatrix} &= C\begin{pmatrix} x \\ y \end{pmatrix} + D\begin{pmatrix} n\xi \\ n\eta \end{pmatrix}. \end{aligned} \tag{II, 2}$$

In Part I we found that between these matrices there are the equations

$$\begin{aligned} \begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} \tilde{D} & -\tilde{B} \\ -\tilde{C} & \tilde{A} \end{pmatrix} &= \begin{pmatrix} E & 0 \\ 0 & E \end{pmatrix}, \\ \begin{pmatrix} \tilde{D} & -\tilde{B} \\ -\tilde{C} & \tilde{A} \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix} &= \begin{pmatrix} E & 0 \\ 0 & E \end{pmatrix}. \end{aligned} \tag{II, 3}$$

The sign $\sim$ means that we have in the particular matrix interchanged lines and columns, for instance:

$$\tilde{D} = \begin{pmatrix} d_{11} & d_{21} \\ d_{12} & d_{22} \end{pmatrix},$$

and E and 0 stand for the unit and zero matrices, respectively:

$$E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad 0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

The Eqs. (II, 3) are identical with the six matrix equations:

$$\begin{aligned} &A\tilde{D} - B\tilde{C} = D\tilde{A} - C\tilde{B} = E, \\ &\tilde{D}A - \tilde{B}C = \tilde{A}D - \tilde{C}B = E, \\ &A\tilde{B} - B\tilde{A} = 0 \qquad C\tilde{D} - D\tilde{C} = 0, \\ &\tilde{D}B - \tilde{B}D = 0 \qquad \tilde{C}A - \tilde{A}C = 0. \end{aligned} \tag{II, 4}$$

We notice that from the first Eqs. (4), we can derive

$$\begin{aligned} A\tilde{D} - D\tilde{A} &= B\tilde{C} - C\tilde{B}, \\ \tilde{A}D - \tilde{D}A &= \tilde{C}B - \tilde{B}C. \end{aligned} \tag{II, 5}$$

Now, we shall investigate how the matrices change if we let the coordinate system rotate by the angle ϕ on the object side and the angle ϕ' on the image side. If we designate the new coordinates by a star, we get

$$\begin{aligned} x^* &= x\cos\phi + y\sin\phi \\ y^* &= x\sin\phi + y\cos\phi, \text{ etc.} \end{aligned} \tag{II, 6}$$

Let us designate the matrix in (6) by $T(\phi)$. We can write

$$\begin{aligned} &\begin{pmatrix} x \\ y \end{pmatrix}^* = T(\phi)\begin{pmatrix} x \\ y \end{pmatrix}, \quad \begin{pmatrix} n\xi \\ n\eta \end{pmatrix}^* = T(\phi)\begin{pmatrix} n\xi \\ n\eta \end{pmatrix}, \\ &\begin{pmatrix} x' \\ y' \end{pmatrix}^* = T(\phi')\begin{pmatrix} x' \\ y' \end{pmatrix}, \quad \begin{pmatrix} n'\xi' \\ n'\eta' \end{pmatrix}^* = T(\phi')\begin{pmatrix} n'\xi' \\ n'\eta' \end{pmatrix}. \end{aligned} \tag{II, 7}$$

From (6) we conclude that

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} &= T(-\phi)\begin{pmatrix} x \\ y \end{pmatrix}^*, \\ \begin{pmatrix} n\xi \\ n\eta \end{pmatrix} &= T(-\phi)\begin{pmatrix} n\xi \\ n\eta \end{pmatrix}^*. \end{aligned} \tag{II, 8}$$

If we introduce (6) and (7) into (2), we find that

$$\begin{aligned} \begin{pmatrix} x' \\ y' \end{pmatrix}^* &= T(\phi')AT(-\phi)\begin{pmatrix} x \\ y \end{pmatrix} \\ &\quad + T(\phi')BT(-\phi)\begin{pmatrix} n\xi \\ n\eta \end{pmatrix}, \\ \begin{pmatrix} n'\xi' \\ n'\eta' \end{pmatrix}^* &= T(\phi')CT(-\phi)\begin{pmatrix} x \\ y \end{pmatrix} \\ &\quad + T(\phi')DT(-\phi)\begin{pmatrix} n\xi \\ n\eta \end{pmatrix}, \end{aligned} \tag{II, 9}$$

or $A^*=T(\phi')AT(-\phi)$, $C^*=T(\phi')CT(-\phi)$,
$B^*=T(\phi')BT(-\phi)$, $D^*=T(\phi')DT(-\phi)$.

The matrices A, B, C, D, are, in general, asymmetric. We have seen in Part I that the symmetric matrices $A\tilde{A}$, $B\tilde{B}$, and $\tilde{A}A$, $\tilde{B}B$, respectively, play an essential role in our investigation. The transformation equations for these matrices can be written

$$\begin{aligned}
(\tilde{A}A)^*&=T(\phi)(\tilde{A}A)T(-\phi),\\
(\tilde{B}B)^*&=T(\phi)(\tilde{B}B)T(-\phi),\\
(\tilde{C}C)^*&=T(\phi)(\tilde{C}C)T(-\phi),\\
(\tilde{D}D)^*&=T(\phi)(\tilde{D}D)T(-\phi),\\
(A\tilde{A})^*&=T(\phi')A\tilde{A}T(\phi'),\\
(B\tilde{B})^*&=T(\phi')B\tilde{B}T(\phi'),\\
(C\tilde{C})^*&=T(\phi')C\tilde{C}T(\phi'),\\
(D\tilde{D})^*&=T(\phi')D\tilde{D}T(\phi').
\end{aligned} \qquad \text{(II, 10)}$$

Eqs. (10) show us that the last four matrices do not vary if we rotate the object coordinate system. If we rotate the image coordinate system, the matrix undergoes a "similarity transformation," which means that it is transformed in the same way as the coefficients of a normal congruence would be. Therefore, we call the last four matrices the image matrices, and the first four, the object matrices.

What can we achieve by rotating the object and image coordinate system? By rotating the object coordinate planes, we can bring one of the first four matrices of (10) into diagonal form (then the coordinate planes coincide with the principal axes of the tensor), and by rotating the image coordinate planes, we can bring one of the last four matrices into diagonal form. If we choose two corresponding matrices, for instance C and C', then C itself becomes a diagonal matrix. The angles ϕ and ϕ' of the rotation necessary for this purpose can easily be calculated. We find that

$$\begin{aligned}
\tan 2\phi&=\frac{2(c_{11}c_{21}+c_{12}c_{22})}{(c_{11}^2+c_{12}^2)-(c_{21}^2+c_{22}^2)},\\
\tan 2\phi'&=\frac{2(c_{11}c_{12}+c_{21}c_{22})}{(c_{11}^2+c_{21}^2)-(c_{12}^2+c_{22}^2)}.
\end{aligned} \qquad \text{(II, 11)}$$

ϕ and ϕ' are undetermined only when C is a so-called pseudo-orthogonal matrix:

$$\begin{aligned}
c_{11}&=\pm c_{22},\\
c_{12}&=\mp c_{21},
\end{aligned} \qquad \text{(II, 12)}$$

or $C\tilde{C}=\tilde{C}C=c_{11}^2E$.

The two matrices are then unchanged by transformation of the coordinate system. Here we can obtain $c_{12}=c_{21}=0$ in an infinite number of ways, if we choose ϕ and ϕ' such that, in case the upper signs in (12) hold,

$$\tan(\phi-\phi')=\frac{c_{21}-c_{12}}{c_{11}+c_{22}}=\frac{c_{21}}{c_{11}},$$

and in the other case (II, 13)

$$\tan(\phi+\phi')=\frac{c_{12}+c_{21}}{c_{11}-c_{22}}=\frac{c_{21}}{c_{11}}.$$

In this case we have still either an object or an image-side rotation at our disposal.

If $C\tilde{C}=0$, C vanishes identically of course, and we have two rotations still at our disposal, so that we can bring one of the other matrices into diagonal form.

Now let us investigate the way the matrices vary if we shift object and image origins along the z, z' axes by the amount z/n, z'/n', respectively. We find that

$$\begin{aligned}
\begin{pmatrix}x'\\y'\end{pmatrix}^*&=\begin{pmatrix}x'\\y'\end{pmatrix}+z'\begin{pmatrix}n'\xi'\\n'\eta'\end{pmatrix},\\
\begin{pmatrix}n'\xi'\\n'\eta'\end{pmatrix}^*&=\begin{pmatrix}n'\xi'\\n'\eta'\end{pmatrix},\\
\begin{pmatrix}x\\y\end{pmatrix}&=\begin{pmatrix}x\\y\end{pmatrix}^*-z\begin{pmatrix}n\xi\\n\eta\end{pmatrix}^*,\\
\begin{pmatrix}n\xi\\n\eta\end{pmatrix}&=\begin{pmatrix}n\xi\\n\eta\end{pmatrix}^*.
\end{aligned} \qquad \text{(II, 14)}$$

Introducing this into (II, 2) we obtain

$$\begin{aligned}
\begin{pmatrix}x'\\y'\end{pmatrix}^*&=(A+z'C)\begin{pmatrix}x\\y\end{pmatrix}^*\\
&\quad+(B+z'D-zA-zz'C)\begin{pmatrix}n\xi\\n\eta\end{pmatrix}^*,\\
\begin{pmatrix}n'\xi'\\n'\eta'\end{pmatrix}^*&=C\begin{pmatrix}x\\y\end{pmatrix}^*+(D-zC)\begin{pmatrix}n\xi\\n\eta\end{pmatrix}^*,
\end{aligned} \qquad \text{(II, 15)}$$

or

$$\begin{aligned}
A^*&=A+z'C, & C^*&=C,\\
B^*&=B-zA+z'D-zz'C, & D^*&=D-zC.
\end{aligned} \qquad \text{(II, 16)}$$

We note that the fundamental equations (II, 4) and (II, 5) of course remain valid for the starred matrices. We note further that C remains unchanged if we shift the origin.

Let us next investigate some matrix functions which we shall soon see are important. If we consider two arbitrary matrices, S and T, we can form the two matrices:

$$\begin{aligned} S\tilde{T}-T\tilde{S}&=K,\\ \tilde{S}T-\tilde{T}S&=L. \end{aligned} \qquad \text{(II, 17)}$$

Both are skew matrices, which means that

$$\begin{aligned} \tilde{K}+K&=0,\\ \tilde{L}+L&=0. \end{aligned} \qquad \text{(II, 18)}$$

That means that we can represent our skew matrices as being proportional to the skew unit matrix, I.

$$I=\begin{pmatrix} 0 & 1\\ -1 & 0 \end{pmatrix}. \qquad \text{(II, 19)}$$

We then have

$$K=\kappa I, \quad L=\lambda I. \qquad \text{(II, 20)}$$

The matrix I has the quality of being invariant against similarity transformations, which means

$$\begin{aligned} T(\phi)IT(-\phi)&=I,\\ T(\phi')IT(-\phi')&=I. \end{aligned} \qquad \text{(II, 21)}$$

Now we can form such combinations with the four matrices A, B, C, D. Some of them vanish when we consider (II, 4 and 5). Let us name the remaining constants α, β, γ, α', β', γ'. We can write

$$\begin{aligned} A\tilde{B}-B\tilde{A}&=0, & \tilde{A}B-\tilde{B}A&=\gamma' I,\\ A\tilde{C}-C\tilde{A}&=\alpha I, & \tilde{A}C-\tilde{C}A&=0,\\ B\tilde{D}-D\tilde{B}&=\gamma I, & \tilde{B}D-\tilde{D}B&=0,\\ C\tilde{D}-D\tilde{C}&=0, & \tilde{C}D-\tilde{D}C&=\alpha' I,\\ A\tilde{D}-D\tilde{A}&=\beta I, & \tilde{A}D-\tilde{D}A&=\beta' I,\\ B\tilde{C}-C\tilde{B}&=\beta I, & \tilde{B}C-\tilde{C}B&=-\beta' I. \end{aligned} \qquad \text{(II, 22)}$$

According to (11, 21), α, β, γ, α', β', γ' do not depend on the positions of the coordinate planes, but only on the locations of the origins. If we apply (II, 16) we find that by a change of origin:

$$\begin{aligned} \alpha^*&=\alpha, & \alpha'^*&=\alpha',\\ \beta^*&=\beta-z\alpha, & \beta'^*&=\beta'-z'\alpha',\\ \gamma^*&=\gamma-2\beta z+z^2\alpha, & \gamma'^*&=\gamma'-2\beta' z'+z'^2\alpha'. \end{aligned} \qquad \text{(II, 23)}$$

We will demonstrate in the fourth section that always

$$\beta^2-\alpha\gamma=\beta'^2-\alpha'\gamma'=\Delta. \qquad \text{(II, 24)}$$

Eqs. (II, 23) show that this magnitude is also invariant against the shifting of the origin and therefore an invariant of the system.

We are now able to classify optical systems according to the sign of Δ. Systems in which $\Delta<0$ we call "torted" systems (Gullstrand: "tordierte Systeme"); those in which $\Delta>0$ we call "retrotorted" systems ("retordierte Systeme"); and those for which $\Delta=0$ we might call "semitorted" systems ("semitordierte Systeme"). The reason for the choice of these names will become clear later on, when we investigate their qualities.

Eqs. (II, 23) show us that if $\alpha(\alpha')$ is not zero, we can choose the origin such that $\beta(\beta')$ vanishes. The point defined in this way we shall call the *main point* (Gullstrand: "Hauptkreuzungspunkt") of the ray. We shall investigate the importance of this point later.

Eqs. (II, 23) shows us that if $\Delta>0$, there exist two points on each ray such that $\gamma(\gamma')$ vanishes. Let us call these *orthogonal points*. This name also will be justified later.

Eqs. (II, 23) shows now that the classifications given above can be defined by the number of real orthogonal points. Retrotorted systems have two real orthogonal points; semitorted systems have one real orthogonal point; and torted systems have none. In addition, there is a fourth class of systems in which each point is an orthogonal point. These we shall call *orthogonal systems*. They are given by

$$\alpha=\beta=\gamma=0, \quad \alpha'=\beta'=\gamma'=0. \qquad \text{(II, 25)}$$

These orthogonal systems will have a special investigation, so we shall continue by assuming that α, β, γ (α', β', γ') do not vanish simultaneously.

Eqs. (II, 23) shows that if $\alpha(\alpha')=0$, the infinite object (image) point is orthogonal. Then certainly we have $\Delta>0$ and a retrotorted system.

This consideration and Eqs. (II, 23) show that we can choose the coordinate system such that

$$\beta=\beta'=0, \qquad \text{(II, 26)}$$

with the one exception of the case where one of

the infinite points is orthogonal ($\alpha(\alpha')=0$). In the retrotorted systems which embrace that particular case, we can always make

$$\gamma=\gamma'=0. \qquad \text{(II, 27)}$$

This will give us some relations between the coefficients of the matrices A, B, C, D. The choice of the origins and the choice of the coordinate planes such that C becomes a diagonal matrix comprise the suggestions the author has to offer toward the standardization of our equations.

Let us investigate the geometrical properties of the orthogonal points before going on to give these standard forms.

III. Orthogonal Points and Fundamental Substitution

We now consider the rays through the object origin. That means that we insert into Eqs. (II, 2) $x=y=0$. We find that

$$\begin{pmatrix}x'\\y'\end{pmatrix}=B\begin{pmatrix}n\xi\\n\eta\end{pmatrix},\quad \begin{pmatrix}n'\xi'\\n'\eta'\end{pmatrix}=D\begin{pmatrix}n\xi\\n\eta\end{pmatrix}, \qquad \text{(III, 1)}$$

or, if the determinant of D does not vanish,

$$\begin{pmatrix}x'\\y'\end{pmatrix}=R\begin{pmatrix}n'\xi'\\n'\eta'\end{pmatrix},\qquad R'=BD^{-1}=\bar{D}^{-1}\bar{B}. \qquad \text{(III, 2)}$$

The tensor of angle magnification is given by

$$\Gamma'=D\bar{D}. \qquad \text{(III, 3)}$$

Generally these two tensors are independent of each other. Let us see what happens if the origin is an orthogonal point. Then we find that

$$B\bar{D}=D\bar{B}. \qquad \text{(III, 4)}$$

Therefore $R'\Gamma'=B\bar{D}=D\bar{B}=\Gamma'R'$,

which means that in an orthogonal point, and only in an orthogonal point, do the two tensors R' and Γ' have the same direction. In this case, we have

$$\begin{matrix}B\bar{D}=D\bar{B},\\ \bar{D}B=\bar{B}D,\end{matrix}\quad\text{or}\quad \begin{matrix}B\bar{B}D\bar{D}=D\bar{D}B\bar{B},\\ \bar{B}B\bar{D}D=\bar{D}D\bar{B}B,\end{matrix} \qquad \text{(III, 5)}$$

which means that the respective tensors have the same direction. We can choose the coordinate system especially so that B and D are diagonal matrices. Then (1) becomes

$$\begin{matrix}x'=b_{11}n\xi, & n'\xi'=d_{11}n\xi,\\ y'=b_{22}n\eta, & n'\eta'=d_{22}n\eta.\end{matrix} \qquad \text{(III, 6)}$$

We see that in this case there exist two orthogonal planes through the principal ray (namely $\xi=0$, $\eta=0$) such that the rays which pass through the origin and lie in these planes are the rays which intersect the axis in the image space. This is the ordinary definition of an orthogonal point. As a rule, the rays in the main sections on the image side lie, on the object side, in two planes which are not orthogonal. Only if these planes are orthogonal do we call the object point an orthogonal point.

That the two definitions we have given are equivalent is easily seen. The origin might be an orthogonal point in the second sense. Then we can choose the coordinate planes according to (6), which gives at once

$$B\bar{D}-D\bar{B}=0 \qquad \text{(III, 7)}$$

and we know, moreover, that this equation is independent of the choice of coordinate planes.

We shall soon learn another equivalent definition and see another interesting feature of the orthogonal points.

Let us now investigate the image congruences which belong to an arbitrary point of the object ray. Eqs. (III, 2) and (II, 16) give us

$$\begin{pmatrix}x'\\y'\end{pmatrix}=R'\begin{pmatrix}n'\xi'\\n'\eta'\end{pmatrix} \qquad \text{(III, 8)}$$

with $$R'(z)=(B-zA)(D-zC)^{-1}=(\bar{D}-z\bar{C})^{-1}(\bar{B}-z\bar{A}).$$

In general, the principal sections of this tensor will rotate around the axis with a change in z. We take a certain point $z=z_1$, and ask if there exist other points (z_2) such that $R'(z_1)$ and $R'(z_2)$ have the same principal sections; or, algebraically, if the matrix equation

$$R'(z_1)R'(z_2)=R'(z_2)R'(z_1) \qquad \text{(III, 9)}$$

has solutions for every z_1, and if so, how many.

We write the matrix equations explicitly

$$(\bar{D}-z_1\bar{C})^{-1}(\bar{B}-z_1\bar{A})(B-z_2A)(D-z_2C)^{-1}=(\bar{D}-z_2\bar{C})^{-1}(\bar{B}-z_2\bar{A})(B-z_1A)(D-z_1C)^{-1}. \qquad \text{(III, 9)}$$

Now we make use of the fact that from equation (II, 22)

$$C\tilde{D}-D\tilde{C}=0$$

we can derive

$$(D-z_1C)(\tilde{D}-z_2\tilde{C})=(D-z_2C)(\tilde{D}-z_1\tilde{C}). \quad \text{(III, 10)}$$

Then we find, multiplying Eq. (III, 9) to right and left by Eq. (III, 10) that

$$(D-z_2C)(\tilde{B}-z_1\tilde{A})(B-z_2A)(\tilde{D}-z_1\tilde{C}) = (D-z_1C)(\tilde{B}-z_2\tilde{A})(B-z_1A)(\tilde{D}-z_2\tilde{C}). \quad \text{(III, 11)}$$

If we apply repeatedly Eq. (II, 22), we find finally from Eq. (III, 11) that

$$z_1z_2\alpha-(z_1+z_2)\beta+\gamma=0 \quad \text{(III, 12)}$$

or

$$z_2=(z_1\beta-\gamma)/(z_1\alpha-\beta).$$

We see that there exists generally for each object point one and only one point such that their main sections on the image side coincide. Eq. (III, 12), therefore, gives a one-to-one coordination of the points of the object ray. We call (III, 12) the object side *fundamental substitution*. We see that the fundamental invariant is the discriminant of the fundamental substitution, and the orthogonal points are nothing but the fixed points of the fundamental substitution, or the points which correspond to themselves according to the fundamental substitution.

We see furthermore that in orthogonal systems *all* points have the same main sections, which means that Eq. (III, 9) is identically fulfilled. What is true for the points of the object ray is, of course, valid in the same way for the points of the image ray. There we find that the points given by $z_1'z_2'$ and

$$z_2'=(z_1'\beta'-\gamma')/(z_1'\alpha'-\beta') \quad \text{(III, 13)}$$

have on the object side the same main section. Eq. (III, 13) gives the fundamental image side substitution.

Now we have a better understanding of the different features of the four types of systems which we distinguished in Section II.

Torted systems

We start with an object point and its main sections on the image side. If we shift the object point, the main sections of the image congruence rotate around the axis. The two points coordinated by (III, 13) move in the same direction without ever coinciding, because the substitution has no real fixed points. Therefore, if we move the object point, the sections of the image side rotate always in one direction around the axis. The only outstanding point is the main point, where the main sections coincide with those of the infinite point.

Retrotorted systems

The main sections of points coordinated by our substitution must coincide. Therefore, as we approach a real orthogonal point, the main sections of the image bundle must have successively different positions, which means that they rotate around the principal ray until we come to the orthogonal point. After we pass this point, the main sections must go through the same positions as before, which means that they must rotate from the orthogonal point on, in the opposite direction. This direction remains unchanged until we come to the second orthogonal point, from which point on the main sections must take up again their original direction of rotation.

Semitorted systems

In these systems there is only one orthogonal point. If we move the object point toward the orthogonal point, the main sections of the image congruence rotate around the axis in one direction until we come to the fixed point. From there they rotate in the opposite direction. The main sections for the fixed point are the same as those for the infinite point. Let us note that in the case of a vanishing determinant according to (III, 13), the fixed point corresponds to each object point, which means that the bundle originating from there has an infinite number of main sections. That is possible only if the fixed point of the semitorted system has a point image, that is that all the rays through the fixed point meet the axis; for in this case and only in this case, is R' the unit matrix which can be interchanged with any other matrix. A point such that all rays proceeding from it meet in a point in the image space shall be called a *stigmatic point*.

We shall now demonstrate analytically that the fixed point of a semitorted system is stigmatic, and the converse, that a stigmatic point in a nonorthogonal system is always the fixed point of a semitorted system. Let us demonstrate the second statement first.

We put the object and image origin at the stigmatic point, and then $x=y=0$ implies $x'=y'=0$. If we insert this into Eq. (II, 2), we see that this implies that

$$B=0. \qquad \text{(III, 14)}$$

Now Eq. (II, 22) shows that

$$\beta=\beta'=\gamma=\gamma'=0, \qquad \text{(III, 15)}$$

which means that the origin is, on both object and image side, the only orthogonal point of a semitorted system. For the proof of the converse, let us put the origins at the only orthogonal point of a semitorted system. Then Eq. (III, 15) is fulfilled, and we have

$$\begin{aligned} A\tilde{B}&=B\tilde{A}, \quad A\tilde{D}=D\tilde{A},\\ \tilde{A}B&=\tilde{B}A, \quad \tilde{A}D=\tilde{D}A,\\ B\tilde{C}&=C\tilde{B}, \quad \tilde{B}C=\tilde{C}B,\\ B\tilde{D}&=D\tilde{B}, \quad \tilde{B}D=\tilde{D}B, \end{aligned} \qquad \text{(III, 16)}$$

and $$A\tilde{C}-C\tilde{A}\neq 0, \quad \tilde{A}C-\tilde{C}A\neq 0.$$

These equations are possible only if $B=0$. Then $x=y=0$ implies $x'=y'=0$, and the object origin is a stigmatic point. Q.E.D.

Let us now give the standard forms for the coefficients of the first-order aberrations and discuss the special cases.

IV. Standard Forms in Nonorthogonal Systems

We learned in Section II that we can choose the origins in a convenient way, and also that we can make a convenient rotation of the coordinate planes such that at least one of the four matrices becomes a diagonal matrix. We shall now show how the sixteen coefficients of A, B, C, D, simplify if we choose convenient origins.

We first let the origins be arbitrary. We begin with four matrices, which we write for convenience

$$\begin{aligned} A&=\begin{pmatrix} a_1+a_2 & a_3+a_4\\ a_3-a_4 & a_1-a_2\end{pmatrix},\\ B&=\begin{pmatrix} b_1+b_2 & b_3+b_4\\ b_3-b_4 & b_1-b_2\end{pmatrix},\\ C&=\begin{pmatrix} c_1+c_2 & c_3+c_4\\ c_3-c_4 & c_1-c_2\end{pmatrix},\\ D&=\begin{pmatrix} d_1+d_2 & d_3+d_4\\ d_3-d_4 & d_1-d_2\end{pmatrix}. \end{aligned} \qquad \text{(IV, 1)}$$

Written in this form, A can be orthogonal only if either $a_1=a_4=0$, or $a_2=a_3=0$. A is diagonal only if $a_3=a_4=0$, and zero only if $a_1=a_2=a_3=a_4=0$, the equivalent holding true for B, C, and D. Between the sixteen coefficients of (IV, 1) there exist the relations (II, 22). If we add and subtract the equations of the same line of (II, 22) we obtain

$$\begin{aligned} \alpha/4&=(a_2c_3-c_2a_3)=(a_4c_1-c_4a_1),\\ \alpha'/4&=(c_2d_3-d_2c_3)=(c_1d_4-(c_1d_4-d_1c_4)),\\ \gamma/4&=(b_2d_3-d_2b_3)=(b_4d_1-d_4b_1),\\ \gamma'/4&=(a_2b_3-b_2a_2)=(a_1b_4-b_1a_4), \end{aligned} \qquad \text{(IV, 2)}$$

and $$\begin{aligned} (\beta+\beta')/4&=a_2d_3-d_2a_3=c_1b_4-b_1c_4,\\ (\beta-\beta')/4&=b_2c_3-b_3c_2=d_1a_4-a_1d_4. \end{aligned}$$

Those are not the only conditions between the coefficients. Closer investigation shows that only five of the six equations are independent, and that we need another equation, derived from $(A\tilde{D}-B\tilde{C}=E)$, for instance,

$$\begin{aligned} &(a_1+a_2)(d_1+d_2)-(b_1+b_2)(c_1+c_2)\\ &+(a_3+a_4)(d_3+d_4)-(b_3+b_4)(c_3+c_4)=1 \end{aligned} \qquad \text{(IV, 3)}$$

in order to eliminate six of the sixteen coefficients. In any case, (IV, 2) gives us the missing proof that the fundamental discriminant is the same on object and image side. We have

$$\begin{aligned} (\beta^2-\beta'^2)/4&=(a_2d_3-d_2a_3)(b_2c_3-c_2b_3)\\ &=(a_2c_3-c_2a_3)(b_2d_3-d_2b_2)\\ &\quad-(c_2d_3-d_2c_3)(b_2d_3-d_2b_3)\\ &=(\alpha\gamma-\alpha'\gamma')/4 \end{aligned} \qquad \text{(IV, 4)}$$

or $$\beta^2-\alpha\gamma=\beta'^2-\alpha'\gamma'.$$

Now, we know from Section II that we can put the origins at the main points, unless one of the orthogonal points is infinite. The latter case

is a special case of retrotorted systems, and in retrotorted systems we can put the origins at orthogonal points.

Thus, we have two standard forms:

A. Origins at main points.
B. Origins at orthogonal points.

We investigate first case *A*. There we have

$$\beta=\beta'=0 \qquad \text{(IV, 5)}$$

or

$$\begin{aligned} &a_2d_3=d_2a_3, \quad b_2c_3=c_2b_3, \quad c_1b_4=b_1c_4, \quad d_1a_4=a_1d_4,\\ &a_2c_3-c_2a_3=a_4c_1-c_4a_1, \quad c_2d_3-d_2c_3=c_1d_4-d_1c_4,\\ &b_2d_3-d_2b_3=b_4d_1-d_4b_1, \quad a_2b_3-b_2a_3=a_1b_4-b_1a_4. \end{aligned} \qquad \text{(IV, 6)}$$

These equations hold true if we set

$$\begin{aligned} &d_1=-\kappa a_1 \quad d_2=\kappa a_2 \quad b_1=\lambda c_1 \quad b_2=-\lambda c_2,\\ &d_4=-\kappa a_4 \quad d_3=\kappa a_3 \quad b_4=\lambda c_4 \quad b_3=-\lambda c_3. \end{aligned} \qquad \text{(IV, 7)}$$

Then (IV, 6) is reduced to the single equation

$$a_2c_3-c_2a_3=a_4c_1-c_4a_1. \qquad \text{(IV, 8)}$$

Let us insert (IV, 7) in (IV, 1). We find for the standard form for the case where object and image origins are at main points

$$\begin{aligned} A&=\begin{pmatrix} a_1+a_2 & a_3+a_4\\ a_3-a_4 & a_1-a_2\end{pmatrix},\\ B&=\lambda\begin{pmatrix} c_1-c_2 & c_4-c_3\\ -c_3-c_4 & c_1+c_2\end{pmatrix},\\ C&=\begin{pmatrix} c_1+c_2 & c_3+c_4\\ c_3-c_4 & c_1-c_2\end{pmatrix},\\ D&=\kappa\begin{pmatrix} a_2-a_1 & a_3-a_4\\ a_3+a_4 & -a_1-a_2\end{pmatrix}. \end{aligned} \qquad \text{(IV, 9)}$$

The ten constants in (IV, 9) are interconnected by the two equations

$$a_2c_3-c_2a_3=a_4c_1-c_4a_1 \qquad \text{(IV, 10)}$$

and

$$\kappa(a_2^2-a_1^2+a_3^2-a_4^2)-\lambda(c_1^2-c_2^2+c_4^2-c_3^2)=1.$$

We find that

$$\begin{aligned} &\alpha/4=a_2c_3-c_2a_3, \qquad \alpha'/4=(a_2c_3-c_2a_3),\\ &\gamma/4=-\kappa\lambda(a_2c_3-c_2a_3), \quad \gamma'/4=-\lambda(a_2c_3-c_2a_3). \end{aligned} \qquad \text{(IV, 11)}$$

The discriminant is positive (negative) if κ and λ have the same (different) sign. In the first case, the system is retrotorted, in the second, torted. The semitorted systems are given by $\lambda=0$. Then we have

$$\begin{aligned} A&=\begin{pmatrix} a_1+a_2 & a_3+a_4\\ a_3-a_4 & a_1-a_2\end{pmatrix}, \quad B=\begin{pmatrix} 0 & 0\\ 0 & 0\end{pmatrix},\\ C&=\begin{pmatrix} c_1+c_2 & c_3+c_4\\ c_3-c_4 & c_1-c_2\end{pmatrix},\\ D&=\kappa\begin{pmatrix} a_2-a_1 & a_3-a_4\\ a_3+a_4 & -a_1-a_2\end{pmatrix} \end{aligned} \qquad \text{(IV, 12)}$$

with

$$a_2c_3-a_3c_2=a_4c_1-a_1c_4, \quad \kappa=1/(a_2^2-a_1^2+a_3^2-a_4^2).$$

By rotating the coordinate system, we can make at least one of the matrices in (IV, 9) a diagonal matrix. In the case of the origins at the main points, we see that *A* and *D*, as well as *B* and *C*, can become diagonal matrices only simultaneously. The reader might investigate these two special cases.

We shall now consider the second case, where the object and image origins are at orthogonal points. This is possible only in retrotorted systems. Here we have $\gamma=\gamma'=0$. Eq. (IV, 4) gives us then $\beta^2=\beta'^2$ or $\beta=\pm\beta'$. If the upper sign holds, we call object and image origins coordinated orthogonal points, and if the lower sign holds, uncoordinated orthogonal points. Let us first assume that the origins are at coordinated orthogonal points. We have

$$\begin{aligned} &\gamma=\gamma'=0\\ &\beta=\beta'\neq 0 \end{aligned} \qquad \text{(IV, 13)}$$

or, according to (IV, 2)

$$\begin{aligned} &b_2d_3-d_3b_2=0, \quad b_4d_1-b_1d_4=0,\\ &b_2a_3-b_3a_2=0, \quad b_4a_1-b_1a_4=0,\\ &d_1a_4-a_1d_4=0,\\ &a_2d_3-a_3d_2\neq 0. \end{aligned} \qquad \text{(IV, 14)}$$

The last equation tells us that the first two allow as the only solution

$$b_2=b_3=0, \qquad \text{(IV, 15)}$$

which means that *B* must be a pseudo-orthogonal

matrix. The Eqs. (IV, 2) then permit the solution

$$\begin{aligned} &b_2=0, \quad a_1=\sigma b_1, \quad d_1=\rho b_1,\\ &b_3=0, \quad a_4=\sigma b_4, \quad d_4=\rho b_4,\\ &c_2=\sigma d_2+\rho a_2,\\ &c_3=\sigma d_3+\rho a_3. \end{aligned} \tag{IV, 16}$$

If we insert this in (IV, 2), we get as the standard form for the matrices

$$A=\begin{pmatrix} a_2+\sigma b_1 & a_3+\sigma b_4 \\ a_3-\sigma b_4 & \sigma b_1-a_2 \end{pmatrix} \quad B=\begin{pmatrix} b_1 & b_4 \\ -b_4 & b_1 \end{pmatrix}, \tag{IV, 17}$$

$$C=\begin{pmatrix} c_1+\sigma d_2+\rho a_2 & c_4+\sigma d_3+\rho a_3 \\ \sigma d_3+\rho a_3-c_4 & c_1-\sigma d_2-\rho a_2 \end{pmatrix},$$

$$D=\begin{pmatrix} \rho b_1+d_2 & d_3+\rho b_4 \\ d_3-\rho b_4 & \rho b_1-d_2 \end{pmatrix}.$$

Between the ten constants of (IV, 17) there are still the two conditions

$$\begin{aligned} &a_2d_3-a_3d_2=c_1b_4-c_4b_1,\\ &(b_1^2+b_4^2)\rho\sigma+a_2d_2+a_3d_3-b_1c_1-b_4c_4=1. \end{aligned} \tag{IV, 18}$$

We find from (IV, 17) that

$$\begin{aligned} &\alpha=\sigma\beta, \quad \alpha'=\rho\beta,\\ &\beta=\beta'=a_2d_3-a_3d_2=c_1b_4-c_4b_1. \end{aligned} \tag{IV, 19}$$

As in the case of the other standard form, we could eliminate two more constants, by the correct choice of coordinate planes. In particular, we could make any of the four matrices A, B, C, D a diagonal matrix.

Eqs. (IV, 19) show us that the object infinite point is orthogonal only if $\sigma=0$, in which case A is pseudo-orthogonal.

$$A=\begin{pmatrix} a_2 & a_3 \\ a_3 & -a_2 \end{pmatrix}, \quad B=\begin{pmatrix} b_1 & b_4 \\ -b_4 & b_1 \end{pmatrix},$$

$$C=\begin{pmatrix} c_1+\rho a_2 & c_4+\rho a_3 \\ \rho a_3-c_4 & c_1-\rho a_2 \end{pmatrix}, \tag{IV, 20}$$

$$D=\begin{pmatrix} \rho b_1+d_2 & d_3+\rho b_4 \\ d_3-\rho b_4 & \rho b_1-d_2 \end{pmatrix}$$

with $\quad a_2d_3-a_3d_2=c_1b_4-c_4b_1,$
$\quad a_2d_2+a_3d_3-c_1b_1-c_4b_4=1.$

The image infinite point is orthogonal if $\rho=0$. In this case D is pseudo-orthogonal, and we ave

$$A=\begin{pmatrix} a_2+\sigma b_1 & a_3+\sigma b_4 \\ a_3-\sigma b_4 & \sigma b_1-a_2 \end{pmatrix}, \quad B=\begin{pmatrix} b_1 & b_4 \\ -b_4 & b_1 \end{pmatrix}, \tag{IV, 21}$$

$$C=\begin{pmatrix} c_1+\sigma d_2 & c_4+\sigma d_3 \\ \sigma d_3-c_4 & c_1-\sigma d_2 \end{pmatrix}, \quad D=\begin{pmatrix} d_2 & d_3 \\ d_3 & -d_2 \end{pmatrix},$$

$$a_2d_3-a_3d_2=c_1b_4-c_4b_1,$$

$$a_2d_2+a_3d_3-c_1b_1-c_4b_4=1.$$

If both infinite points are orthogonal, we have $\rho=\sigma=0$, and all four matrices become pseudo-orthogonal:

$$A=\begin{pmatrix} a_2 & a_3 \\ a_3 & -a_2 \end{pmatrix}, \quad C=\begin{pmatrix} c_1 & c_4 \\ -c_4 & c_1 \end{pmatrix}, \tag{IV, 22}$$

$$B=\begin{pmatrix} b_1 & b_4 \\ -b_4 & b_1 \end{pmatrix}, \quad D=\begin{pmatrix} d_2 & d_3 \\ d_3 & -d_2 \end{pmatrix}$$

with $\quad a_2d_3-a_3d_2=c_1b_4-c_4b_1,$
$\quad a_2d_2+a_3d_3-c_1b_1-c_4b_4=1.$

In the same way, we can deal with the case where the origins are at uncoordinated orthogonal points. We have

$$\begin{aligned} &\gamma=\gamma'=0,\\ &\beta+\beta'=0. \end{aligned} \tag{IV, 23}$$

Eq. (IV, 2) gives, after an equivalent calculation

$$\begin{aligned} &b_1=b_4=0, \quad d_2=\rho b_2, \quad a_2=\sigma b_2,\\ &\qquad\quad d_3=\rho b_3, \quad a_3=\sigma b_3,\\ &c_1=\sigma d_1+\rho a_1,\\ &c_4=\sigma d_4+\rho a_4, \end{aligned} \tag{IV, 24}$$

so that we find for the standard form

$$A=\begin{pmatrix} a_1+\sigma b_2 & a_4+\sigma b_3 \\ \sigma b_3-a_4 & a_1-\sigma b_2 \end{pmatrix}, \quad B=\begin{pmatrix} b_2 & b_3 \\ b_3 & -b_2 \end{pmatrix},$$

$$C=\begin{pmatrix} c_2+\sigma d_1+\rho a_1 & c_3+\sigma d_4+\rho a_4 \\ c_3-\sigma d_4-\rho a_4 & \sigma d_1+\rho a_1-c_2 \end{pmatrix}, \tag{IV, 25}$$

$$D=\begin{pmatrix} d_1+\rho b_2 & d_4+\rho b_3 \\ \rho b_3-d_4 & d_1-\rho b_2 \end{pmatrix}$$

with $\quad b_2c_3-c_2b_3=d_1a_4-a_1d_4,$
$\quad \sigma\rho(b_2^2+b_3^2)+a_1d_1+a_4d_4-c_2b_2-c_3b_3=1$

and $\quad \alpha/4=\sigma(b_2c_3-c_2b_3), \quad \alpha'/4=\rho(c_2b_3-c_3b_2).$
$\quad \beta/4=-\beta'/4=\tfrac{1}{2}(b_2c_3-c_2b_3).$

We see again that the object infinite point is orthogonal if $\sigma=0$, the image infinite point, if $\rho=0$. We give the formulae for the case when both infinite points are orthogonal

$$A=\begin{pmatrix} a_1 & a_4 \\ -a_4 & a_1 \end{pmatrix},\quad B=\begin{pmatrix} b_2 & b_3 \\ b_3 & -b_2 \end{pmatrix},$$
$$C=\begin{pmatrix} c_2 & c_3 \\ c_3 & -c_2 \end{pmatrix},\quad D=\begin{pmatrix} d_1 & d_4 \\ -d_4 & d_1 \end{pmatrix} \qquad \text{(IV, 26)}$$

to compare with (IV, 22).

Next, let us investigate specially the semitorted systems.

V. Semitorted Systems

Let us assume first that the only orthogonal point of the system is finite in object and image space. We put the origin at this point and find that

$$\alpha\neq 0,\quad \alpha'\neq 0,$$
$$\beta=\beta'=0, \qquad \text{(V, 1)}$$
$$\gamma=\gamma'=0,$$

which gives
$$b_1=b_2=b_3=b_4=0,$$
$$d_1=\kappa a_1,\qquad d_2=-\kappa a_2, \qquad \text{(V, 2)}$$
$$d_4=\kappa a_4,\qquad d_3=-\kappa a_3$$

or
$$A=\begin{pmatrix} a_1+a_2 & a_3+a_4 \\ a_3-a_4 & a_1-a_2 \end{pmatrix},\quad B=\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$
$$C=\begin{pmatrix} c_1+c_2 & c_3+c_4 \\ c_3-c_4 & c_1-c_2 \end{pmatrix}, \qquad \text{(V, 3)}$$
$$D=\kappa\begin{pmatrix} a_1-a_2 & a_4-a_3 \\ -a_3-a_4 & a_1+a_2 \end{pmatrix}$$

with
$$\kappa(a_1^2-a_2^2+a_4^2-a_3^2)=1.$$

With the right choice of coordinate planes, we can get either $a_3=a_4=0$, which means

$$A=\begin{pmatrix} a_1+a_2 & 0 \\ 0 & a_1-a_2 \end{pmatrix},\quad B=\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$
$$C=\begin{pmatrix} c_1+c_2 & c_3+c_4 \\ c_3-c_4 & c_1-c_2 \end{pmatrix}, \qquad \text{(V, 4)}$$
$$D=\kappa\begin{pmatrix} a_1-a_2 & 0 \\ 0 & a_1+a_2 \end{pmatrix},$$
$$\kappa(a_1^2-a_2^2)=1,$$

or we can make C a diagonal matrix.

We shall investigate (V, 4) in greater detail. We see that $a_1\neq\pm a_2$, because κ would otherwise become infinite.

If we choose the coordinate planes according to (V, 4), the Eqs. (I, 1) become

$$x'=(a_1+a_2)x,$$
$$y'=(a_1-a_2)y, \qquad \text{(V, 5)}$$
$$n'\xi'=(c_1+c_2)x+(c_3+c_4)y+\frac{1}{a_1+a_2}n\xi,$$
$$n'\eta'=(c_3-c_4)x+(c_1-c_2)y+\frac{1}{a_1-a_2}n\eta.$$

The first two equations show us that in this case there corresponds to each point of the object plane (x, y) a point (x', y') of the image plane such that every ray through the object point passes through the image point; that is, we have a sharp image of the object plane. However, this image is generally not similar, because the magnification in one direction is a_1+a_2, and in the other a_1-a_2. Therefore, we see that a surface element at the stigmatic point of a semitorted system has a sharp but dissimilar image.

Let us now investigate a system where the object infinite point is the only stigmatic point of a semitorted system. We place the object origin arbitrarily, and the image origin at the finite image point, or *focal point*, and find that

$$\alpha=0,\quad \alpha'\neq 0,$$
$$\beta=0,\quad \beta'=0, \qquad \text{(V, 6)}$$
$$\gamma\neq 0,\quad \gamma'=0.$$

This gives, according to (IV, 2)

$$a_1=a_2=a_3=a_4=0,$$
$$b_1=\kappa c_1,\quad b_2=-\kappa c_2, \qquad \text{(V, 7)}$$
$$b_4=\kappa c_4,\quad b_3=-\kappa c_3,$$

or
$$A=\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},\quad B=\kappa\begin{pmatrix} c_1-c_2 & c_4-c_3 \\ -c_3-c_4 & c_1+c_2 \end{pmatrix},$$
$$C=\begin{pmatrix} c_1+c_2 & c_3+c_4 \\ c_3-c_4 & c_1-c_2 \end{pmatrix}, \qquad \text{(V, 8)}$$
$$D=\begin{pmatrix} d_1+d_2 & d_3+d_4 \\ d_3-d_4 & d_1-d_2 \end{pmatrix},$$
$$-\kappa(c_1^2-c_2^2)=1.$$

By choosing the right coordinate planes, we can make either D or B, as well as C, a diagonal matrix. In the latter case, we get

$$A=\begin{pmatrix}0&0\\0&0\end{pmatrix},\quad B=\kappa\begin{pmatrix}c_1-c_2&0\\0&c_1+c_2\end{pmatrix},$$
$$C=\begin{pmatrix}c_1+c_2&0\\0&c_1-c_2\end{pmatrix},\qquad (V,\ 9)$$
$$D=\begin{pmatrix}d_1+d_2&d_3+d_4\\d_3-d_4&d_1-d_2\end{pmatrix}.$$

The degree of freedom that is left (the position of the object origin) can be utilized in several different ways, for instance, to make $c_1=0$.

If the image infinite point is orthogonal, we have

$$\alpha'=\beta=\beta'=\gamma'=0,\quad \alpha\neq0,\quad \gamma\neq0 \qquad (V,\ 10)$$

and we find in the same way that

$$A=\begin{pmatrix}a_1+a_2&a_3+a_4\\a_3-a_4&a_1-a_2\end{pmatrix},$$
$$B=\kappa\begin{pmatrix}c_1-c_2&c_4-c_3\\-c_3-c_4&c_1+c_2\end{pmatrix},$$
$$C=\begin{pmatrix}c_1+c_2&c_3+c_4\\c_3-c_4&c_1-c_2\end{pmatrix},\quad D=\begin{pmatrix}0&0\\0&0\end{pmatrix}, \qquad (V,\ 11)$$
$$\kappa(c_1^2-c_2^2+c_4^2-c_3^2)=1,$$

if we choose the origin on the object side at the focal point and leave the image origin arbitrary.

Finally, we investigated the afocal systems. These are given by

$$\alpha=\alpha'=0,\quad \gamma\neq0,\quad \beta=\beta'=0,\quad \gamma'\neq0. \qquad (V,\ 12)$$

Object and image origins are arbitrary. We find that

$$A=\begin{pmatrix}a_1+a_2&a_3+a_4\\a_3-a_4&a_1-a_2\end{pmatrix},$$
$$B=\begin{pmatrix}b_1+b_2&b_3+b_4\\b_3-b_4&b_1-b_2\end{pmatrix},$$
$$C=\begin{pmatrix}0&0\\0&0\end{pmatrix},\quad D=\kappa\begin{pmatrix}a_1-a_2&a_4-a_3\\-a_3-a_4&a_1+a_2\end{pmatrix}, \qquad (V,\ 13)$$
$$\kappa(a_1^2-a_2^2+a_4^2-a_3^2)=1.$$

And the choice of coordinate planes bring us

$$A=\begin{pmatrix}a_1+a_2&0\\0&a_1-a_2\end{pmatrix},$$
$$B=\begin{pmatrix}b_1+b_2&b_3+b_4\\b_3-b_4&b_1-b_2\end{pmatrix},\quad C=\begin{pmatrix}0&0\\0&0\end{pmatrix}, \qquad (V,\ 14)$$
$$D=\frac{1}{a_1^2-a_2^2}\begin{pmatrix}a_1-a_2&0\\0&a_1+a_2\end{pmatrix}.$$

VI. The Orthogonal Systems

In orthogonal systems, we have

$$\alpha=\alpha'=\beta=\beta'=\gamma=\gamma'=0, \qquad (VI,\ 1)$$

Inserting this in (IV, 2), we find the solutions

$$\begin{aligned}a_3&=\kappa a_2,\quad a_4=\lambda a_1,\\ b_3&=\kappa b_2,\quad b_4=\lambda b_1,\\ c_3&=\kappa c_2,\quad c_4=\lambda c_1,\\ d_3&=\kappa d_2,\quad d_4=\lambda d_1,\end{aligned} \qquad (VI,\ 2)$$

$$A=\begin{pmatrix}a_1+a_2&\kappa a_2+\lambda a_1\\\kappa a_2-\lambda a_1&a_1-a_2\end{pmatrix},$$
$$B=\begin{pmatrix}b_1+b_2&\kappa b_2+\lambda b_1\\\kappa b_2-\lambda b_1&b_1-b_2\end{pmatrix},$$
$$C=\begin{pmatrix}c_1+c_2&\kappa c_2+\lambda c_1\\\kappa c_2-\lambda c_1&c_1-c_2\end{pmatrix}, \qquad (VI,\ 3)$$
$$D=\begin{pmatrix}d_1+d_2&\kappa d_2+\lambda d_1\\\kappa d_2-\lambda d_1&d_1-d_2\end{pmatrix}.$$

We see that if we make one matrix a diagonal matrix, it implies that

$$\kappa=\lambda=0,$$

and then all the matrices assume a diagonal form. If we choose these planes, we can write

$$A=\begin{pmatrix}a_1+a_2&0\\0&a_1-a_2\end{pmatrix},\quad B=\begin{pmatrix}b_1+b_2&b\\0&b_1-b_2\end{pmatrix}, \qquad (VI,\ 4)$$
$$C=\begin{pmatrix}c_1+c_2&0\\0&c_1-c_2\end{pmatrix},\quad D=\begin{pmatrix}d_1+d_2&0\\0&d_1-d_2\end{pmatrix},$$
$$A\tilde{D}-B\tilde{C}=E.$$

Let us ask now if there are any stigmatic points. If there are, we must be able to put object

and image origin at these points and obtain

$$B^*=B-zA+z'D-zz'C=0.$$

This gives two bilinear equations for z and z':

$$\begin{aligned} b_1-za_1+z'd_1-zz'c_1&=0,\\ b_2-za_2+z'd_2-zz'c_2&=0. \end{aligned} \qquad \text{(VI, 5)}$$

Elimination of z and z', respectively, leads to the equations

$$\begin{aligned} &z'=(b_1-za_1)/(zc_1-d_1)=(b_2-za_2)/(zc_2-d_2),\\ &z=(b_1+z'd_1)/(a_1+z'c_1)=(b_2+z'd_2)/(a_2+z'c_2), \end{aligned} \qquad \text{(VI, 6)}$$

or

$$\begin{aligned} &z^2(a_1c_2-a_2c_1)-z[(b_1c_2-b_2c_1)\\ &\quad+(a_1d_2-a_2d_1)]+(b_1d_2-b_2d_1)=0,\\ &z'^2(c_1d_2-c_2d_1)-z'[(b_1c_2-b_2c_1)\\ &\quad+(a_1d_2-a_2d_1)]-(a_1b_2-a_2b_1)=0. \end{aligned} \qquad \text{(VI, 7)}$$

We can write these

$$\begin{aligned} z^2\mathbf{A}-2z\mathbf{B}+\mathbf{C}&=0,\\ z'^2\mathbf{A}'-2z'\mathbf{B}'+\mathbf{C}'&=0, \end{aligned} \qquad \text{(VI, 7)}$$

where

$$\begin{aligned} &\mathbf{A}=a_1c_2-a_2c_1, \quad \mathbf{A}'=c_1d_2-c_2d_1,\\ &\mathbf{C}=b_1d_2-b_2d_1, \quad \mathbf{C}'=a_1v_2-a_2b_1,\\ &\mathbf{B}+\mathbf{B}'=b_1c_2-c_1b_2,\\ &\mathbf{B}-\mathbf{B}'=a_1d_2-d_1a_2. \end{aligned} \qquad \text{(VI, 8)}$$

We notice that the two Eqs. (VI, 7) have the same discriminant D, for from

$$B^2-B'^2=AC-A'C'$$

It follows that $D=B^2-AC=B'^2-A'C'$.

We are, therefore, again able to make a classification according to this quantity, dividing the orthogonal systems into four classes: (1) *Duostigmatic* systems, which have two stigmatic points; (2) *Monostigmatic* systems, which have one stigmatic point; (3) *Full Astigmatic* systems, which have no real stigmatic point; and (4) *Gaussian* systems, in which each point is stigmatic.

We have again found suitable places for the origins.

At any time, unless one of the infinite points is a stigmatic point, we can place the origins so that the coefficients of z and z' in (7) vanish. We call these points the *middle-points* of the ray. At these points $B=B'=0$.

In duostigmatic systems (and to those belong the systems excepted above) we can put the origins at noncoordinated stigmatic points ($C=C'=0$).

First standard form

Origins at middle points. Eq. (VI, 8) gives

$$b_1c_2=b_2c_1, \quad a_1d_2=a_2d_1, \qquad \text{(VI, 9)}$$

or

$$\frac{a_1}{a_2}=\frac{d_1}{d_2}=\kappa, \quad \frac{b_1}{b_2}=\frac{c_1}{c_2}=\lambda, \qquad \text{(VI, 10)}$$

which means that

$$\begin{aligned} A&=\begin{pmatrix} a & 0\\ 0 & \kappa a\end{pmatrix}, \quad B=\begin{pmatrix} b & 0\\ 0 & \lambda b\end{pmatrix},\\ C&=\begin{pmatrix} c & 0\\ 0 & \lambda c\end{pmatrix}, \quad D=\begin{pmatrix} d & 0\\ 0 & \kappa d\end{pmatrix}, \end{aligned} \qquad \text{(VI, 11)}$$

with

$$ad=\frac{\lambda^2-1}{\lambda^2-\kappa^2}, \quad bc=\frac{\kappa^2-1}{\lambda^2-\kappa^2}.$$

The equation for the stigmatic points degenerates to

$$z^2ac=bd, \quad z'^2cd=ab. \qquad \text{(VI, 12)}$$

The system is duostigmatic if $abcd>0$, full astigmatic if $abcd<0$.

Second standard form

Origin at nonconjugated stigmatic points. We have

$$b_1d_2-b_2d_1=0, \quad a_1b_2-a_2b_1=0, \qquad \text{(VI, 13)}$$

or

$$\frac{a_1}{a_2}=\frac{b_1}{b_2}=\frac{d_1}{d_2}=\kappa.$$

Therefore we find that

$$\begin{aligned} A&=\begin{pmatrix} a & 0\\ 0 & \kappa a\end{pmatrix}, \quad B=\begin{pmatrix} b & 0\\ 0 & \kappa b\end{pmatrix},\\ C&=\begin{pmatrix} c & 0\\ 0 & \lambda c\end{pmatrix}, \quad D=\begin{pmatrix} d & 0\\ 0 & \kappa d\end{pmatrix} \end{aligned} \qquad \text{(VI, 14)}$$

with

$$ad=\frac{\kappa\lambda-1}{\kappa\lambda-\kappa^2}, \quad bc=\frac{\kappa^2-1}{\kappa\lambda-\kappa^2}.$$

Eq. (VI, 6) becomes

$$z^2a=zb, \quad z'^2d=zc. \qquad \text{(VI, 15)}$$

The secondary standard form give us at once a very important law. We have

$$\begin{aligned} x' &= ax \quad + bn\xi, \\ y' &= \kappa ay \quad + \kappa bn\eta, \\ n'\xi' &= cx \quad + dn\xi, \\ n'\eta' &= \lambda cy \quad + \kappa dn\eta. \end{aligned} \tag{VI, 16}$$

We consider the rays through the object origin ($x=y=0$). We find from (VI, 11) that

$$\begin{aligned} n'\xi'_0 &= dn\xi_0, \\ n'\eta'_0 &= \kappa dn\eta_0. \end{aligned} \tag{VI, 17}$$

The angle magnification (and therefore the magnification with which a surface element perpendicular to the axis is imaged) is different in different directions. The value κ is a measure of the *dissimilarity*.

Let us now consider the rays through the image origin, that is, the rays through the other stigmatic point ($x'=y'=0$). Elimination of x and y gives

$$\begin{aligned} an'\xi'^0 &= n\xi, \\ \kappa an'\eta'^0 &= n\eta. \end{aligned} \tag{VI, 18}$$

We see that *the dissimilarity (in magnification) at the two stigmatic points has reciprocal values* (Abbe's Law). Only in monostigmatic systems can we have a similar image of an object surface element.

We shall now, analogously to the preceding section, give standard forms for special cases.

(a) One stigmatic point on object side infinite. $A=C=C'=0$,

$$a=0, \quad \lambda=1/\kappa,$$

$$A=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}, \qquad B=-b\begin{pmatrix}1 & 0\\ 0 & \kappa\end{pmatrix},$$

$$C=-\frac{1}{b\kappa}\begin{pmatrix}\kappa & 0\\ 0 & 1\end{pmatrix}, \quad D=d\begin{pmatrix}1 & 0\\ 0 & \kappa\end{pmatrix} \tag{VI, 19}$$

or

$$\begin{aligned} x' &= \quad bn\xi, \\ y' &= \quad \kappa bn\eta, \\ n'\xi' &= -(1/b)x \quad + dn\xi, \\ n'\eta' &= -(1/\kappa b)y + \kappa dn\eta. \end{aligned}$$

(b) One stigmatic point on image side infinite. ($A'=C=C'=0$)

$$A=a\begin{pmatrix}1 & 0\\ 0 & \kappa\end{pmatrix}, \qquad B=b\begin{pmatrix}1 & 0\\ 0 & \kappa\end{pmatrix},$$

$$C=-\frac{1}{\kappa b}\begin{pmatrix}\kappa & 0\\ 0 & 1\end{pmatrix}, \quad D=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}, \tag{VI, 20}$$

or

$$\begin{aligned} x' &= ax \quad + bn\xi, \\ y' &= \quad \kappa ay \quad + \kappa bn\eta, \\ n'\xi' &= -(1/b)x, \\ n'\eta' &= -(1/\kappa b)y. \end{aligned}$$

(c) Both infinite points stigmatic ($A=A'=C=C'=0$).

$$A=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}, \qquad B=b\begin{pmatrix}1 & 0\\ 0 & \kappa\end{pmatrix},$$

$$C=-\frac{1}{\kappa b}\begin{pmatrix}\kappa & 0\\ 0 & 1\end{pmatrix}, \quad D=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}, \tag{VI, 21}$$

or

$$\begin{aligned} x' &= bn\xi, \\ y' &= \quad \kappa bn\eta, \\ n'\xi' &= -(1/b)n\xi, \\ n'\eta' &= \quad -(1/\kappa b)n\eta. \end{aligned}$$

In the same way, let us now make charts for the monostigmatic systems.

(a) Stigmatic points finite on both sides ($B=B'=C=C'=0$).

$$b=0, \quad \kappa=1,$$

$$A=a\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \quad B=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix},$$

$$C=c\begin{pmatrix}1 & 0\\ 0 & \lambda\end{pmatrix}, \quad D=\frac{1}{a}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \tag{VI, 22}$$

or

$$\begin{aligned} x' &= ax, \\ y' &= \quad ay, \\ n'\xi' &= cx \quad + (1/a)n\xi, \\ n'\eta' &= \lambda cy \quad + (1/a)n\eta. \end{aligned}$$

We see that we have a similar image of a surface element at the object point.

(b) Monostigmatic point infinite on object side ($A=B=B'=C'=0$).

$$A=\begin{pmatrix}0&0\\0&0\end{pmatrix},\qquad B=b\begin{pmatrix}1&0\\0&1\end{pmatrix},$$
$$C=-\frac{1}{c}\begin{pmatrix}1&0\\0&1\end{pmatrix},\quad D=d\begin{pmatrix}1&0\\0&\kappa\end{pmatrix},\tag{VI, 23}$$

or
$$\begin{aligned}x'&=&&bn\xi,\\ y'&=&&bn\eta,\\ n'\xi'&=-(1/c)x&&+dn\xi,\\ n'\eta'&=&-(1/c)y&+\kappa dn\eta.\end{aligned}$$

(c) Monostigmatic point infinite on image side ($A'=B=B'=C=0$).

$$A=a\begin{pmatrix}1&0\\0&\kappa\end{pmatrix},\qquad B=b\begin{pmatrix}1&0\\0&1\end{pmatrix},$$
$$C=-\frac{1}{c}\begin{pmatrix}1&0\\0&1\end{pmatrix},\quad D=\begin{pmatrix}0&0\\0&0\end{pmatrix}.\tag{VI, 24}$$

(d) Afocal system monostigmatic ($A=A'=C=C'=0$).

$$A=a\begin{pmatrix}1&0\\0&1\end{pmatrix},\quad B=b\begin{pmatrix}1&0\\0&\kappa\end{pmatrix},$$
$$C=\begin{pmatrix}0&0\\0&0\end{pmatrix},\quad D=\begin{pmatrix}1&0\\0&1\end{pmatrix},\tag{VI, 25}$$

or
$$\begin{aligned}x'&=ax&&+bn\xi,\\ y'&=&ay&+\kappa bn\eta,\\ n'\xi'&=&&dn\xi,\\ n'\eta'&=&&dn\eta.\end{aligned}$$

(e) Gaussian systems.

Finally, we must deal with the systems where each point is stigmatic. This leads to $\kappa=\lambda=1$,

$$A=a\begin{pmatrix}1&0\\0&1\end{pmatrix},\quad B=b\begin{pmatrix}1&0\\0&1\end{pmatrix},$$
$$C=c\begin{pmatrix}1&0\\0&1\end{pmatrix},\quad D=d\begin{pmatrix}1&0\\0&1\end{pmatrix},\tag{VI, 26}$$

with $ad-bc=1$.

Here we can find for each object point the stigmatic image point by making

$$\begin{aligned}B^*&=B-zA+z'D-zz'C\\ &=(b-za+z'd-zz'c)E=0.\end{aligned}\tag{VI, 27}$$

This gives the image point for each object point. Then where is it best to put the origin? According to (II, 16), we have

$$\begin{aligned}a^*&=a+z'c,\\ b^*&=b-za+z'd-zz'c,\\ c^*&=c,\\ d^*&=d-zc.\end{aligned}\tag{VI, 28}$$

We must distinguish two cases. First, $c\neq 0$. Then, we can choose the origin such that a and d vanish. We find for the amount of shift that

$$\begin{aligned}z'&=-(a/c),\\ z&=(d/c).\end{aligned}\tag{VI, 29}$$

These points we call the focal points of the system. If we put the origins there, we get

$$A=\begin{pmatrix}0&0\\0&0\end{pmatrix},\quad B=\frac{1}{c}\begin{pmatrix}1&0\\0&-1\end{pmatrix},$$
$$C=c\begin{pmatrix}1&0\\0&1\end{pmatrix},\quad D=\begin{pmatrix}0&0\\0&0\end{pmatrix},\tag{VI, 30}$$

and our standardized equations become

$$\begin{aligned}x'&=-(1/c)n\xi,\\ y'&=&-(1/c)n\eta,\\ n'\xi'&=cx,\\ n'\eta'&=cy,\end{aligned}\tag{VI, 31}$$

which are the already well-known focal equations for this system.

The second case is given by $c=0$. These are the afocal systems. Instead of (VI, 25) we get

$$\begin{aligned}a^*&=a,\\ b^*&=b-za+z'd,\\ c^*&=0,\\ d^*&=d=1/a.\end{aligned}\tag{VI, 32}$$

We see that a and d are unchanged, but we can put the origins at conjugate points, and so make b equal to zero. Then we have

$$A=a\begin{pmatrix}1&0\\0&1\end{pmatrix},\quad B=\begin{pmatrix}0&0\\0&0\end{pmatrix},$$
$$C=\begin{pmatrix}0&0\\0&0\end{pmatrix},\quad D=\frac{1}{a}\begin{pmatrix}1&0\\0&1\end{pmatrix},\tag{VI, 33}$$

or
$$\begin{aligned}x'&=ax,\\ y'&=&ay,\\ n'\xi'&=&&(1/a)n\xi,\\ n'\eta'&=&&&(1/a)n\eta.\end{aligned}$$

Eqs. (VI, 33) contain all the laws of afocal systems.

VII. Special Problems

We saw in section IV that B is pseudo-orthogonal if we put the origins at orthogonal points. But the converse is true, too. Because

$$b_2=b_3=0 \qquad \text{(VII, 1)}$$

demands, according to (IV, 3), that

$$\gamma=\gamma'=0, \quad \beta=\beta',$$

and $\quad b_1=b_4=0$

requires that $\gamma=\gamma'=0, \quad \beta=-\beta'.$ (VII, 2)

B is zero only if

$$\begin{aligned} &b_1=b_2=b_3=b_4=0, \\ \text{or} \quad &\beta=\beta'=\gamma=\gamma'=0, \end{aligned} \qquad \text{(VII, 3)}$$

which means that object and image origins are conjugated stigmatic points.

In accord with the tetrality principle, we find analogously that A is pseudo-orthogonal if, and only if, the infinite object point and the image origin are orthogonal; and zero if, and only if, the infinite object point is stigmatic and the image origin the conjugated image of this point. C is pseudo-orthogonal if, and only if, both infinite points are orthogonal; and zero in an afocal system, where the two infinite points are conjugated stigmatic points. D is pseudo-orthogonal if, and only if, object origin and image infinite point are orthogonal; zero, if they are conjugated stigmatic points.

Let us finally give some analogies to the well-known laws of Gaussian optics.

If $C=0$, we have

$$\begin{aligned} \begin{pmatrix} x' \\ y' \end{pmatrix} &= A\begin{pmatrix} x \\ y \end{pmatrix} + B\begin{pmatrix} n\xi \\ n\eta \end{pmatrix}, \\ \begin{pmatrix} n'\xi' \\ n'\eta' \end{pmatrix} &= D\begin{pmatrix} n\xi \\ n\eta \end{pmatrix}. \end{aligned} \qquad \text{(VII, 4)}$$

We see that for all object points there is the same tensor of angle magnification:

$$\tilde{D}D=\Gamma_0 \qquad \text{(VII, 5)}$$

and the same is true of the image points because

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} &= \tilde{D}\begin{pmatrix} x' \\ y' \end{pmatrix} - \tilde{B}\begin{pmatrix} n'\xi' \\ n'\eta' \end{pmatrix}, \\ \begin{pmatrix} n\xi \\ n\eta \end{pmatrix} &= \tilde{A}\begin{pmatrix} n'\xi' \\ n'\eta' \end{pmatrix}, \end{aligned} \qquad \text{(VII, 6)}$$

which gives $\quad \Gamma'=\tilde{A}A=(D\tilde{D})^{-1}, \qquad \text{(VII, 7)}$

since $\quad A\tilde{D}=E. \qquad \text{(VII, 8)}$

Now, if $B=0$, we find that

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = A\begin{pmatrix} x \\ y \end{pmatrix}, \quad \begin{pmatrix} x \\ y \end{pmatrix} = \tilde{D}\begin{pmatrix} x' \\ y' \end{pmatrix},$$

and $$\begin{pmatrix} n'\xi' \\ n'\eta' \end{pmatrix} = C\begin{pmatrix} x \\ y \end{pmatrix} + D\begin{pmatrix} n\xi \\ n\eta \end{pmatrix}, \qquad \text{(VII, 9)}$$

$$\begin{pmatrix} n\xi \\ n\eta \end{pmatrix} = -\tilde{C}\begin{pmatrix} x' \\ y' \end{pmatrix} + \tilde{A}\begin{pmatrix} n'\xi' \\ n'\eta' \end{pmatrix}.$$

We see that the projection tensor $\tilde{A}A$ with which the coordinate planes are projected on each other is independent of the projection center, or in other words, the system gives a sharp (but in general not similar) image of the object on the image coordinate plane. If $A=0$, we find that

$$\begin{aligned} &\begin{pmatrix} x' \\ y' \end{pmatrix} = B\begin{pmatrix} n\xi \\ n\eta \end{pmatrix}, \quad \begin{pmatrix} x \\ y \end{pmatrix} = \tilde{D}\begin{pmatrix} x' \\ y' \end{pmatrix} - \tilde{B}\begin{pmatrix} n'\xi' \\ n'\eta' \end{pmatrix} \\ &\begin{pmatrix} n'\xi' \\ n'\eta' \end{pmatrix} = C\begin{pmatrix} x \\ y \end{pmatrix} + D\begin{pmatrix} n\xi \\ n\eta \end{pmatrix}, \quad \begin{pmatrix} n\xi \\ n\eta \end{pmatrix} = -\tilde{C}\begin{pmatrix} x' \\ y' \end{pmatrix} \end{aligned} \qquad \text{(VII, 10)}$$

with $\quad -B\tilde{C}=E.$

In this case, there is a unique focal point on the image side because the infinite object point is stigmatic. We see that the image coordinate plane has the same apparent distance from any object point given by the tensor $\tilde{B}B$.

On the other hand, if $D=0$, we get

$$\begin{aligned} &\begin{pmatrix} x' \\ y' \end{pmatrix} = A\begin{pmatrix} x \\ y \end{pmatrix} + B\begin{pmatrix} n\xi \\ n\eta \end{pmatrix}, \quad \begin{pmatrix} x \\ y \end{pmatrix} = -\tilde{B}\begin{pmatrix} n'\xi' \\ n'\eta' \end{pmatrix}, \\ &\begin{pmatrix} n'\xi' \\ n'\eta' \end{pmatrix} = C\begin{pmatrix} x \\ y \end{pmatrix}, \quad \begin{pmatrix} n\xi \\ n\eta \end{pmatrix} = -\tilde{C}\begin{pmatrix} x' \\ y' \end{pmatrix} + \tilde{A}\begin{pmatrix} n'\xi' \\ n'\eta' \end{pmatrix}, \end{aligned} \qquad \text{(VII, 11)}$$

which means that if the image infinite point is stigmatic, the focal plane in the object space

seems to be the same distance from every image point.

These are qualities which depend on the fact that a finite or infinite point is a stigmatic point; their general validity in Gaussian optics comes only from the fact that there every point is a stigmatic point.

VIII. The Realization of the General Optical System

We shall now show how to find A, B, C, D, if we must calculate a ray and its neighbors through an optical system. As in the corresponding section of the work on Gaussian optics,[2] we have two steps. First, we give the transformation formulae for the refraction at a surface, and then the transfer formulae for passing from surface to surface.

Let us attack the first problem. We take as origin the vertex where the ray meets the refracting surface, and as one coordinate plane, the plane which contains the incident ray and the surface normal (and therefore the refracted ray). The other planes are determined by this choice, since they must be perpendicular to the meridian plane, and the z axis must be the incident ray, the z' axis the emergent ray.

Now, in general, this meridian plane need not correspond with one of the main sections of the refracting surface. It might form with it an angle ω.

We call ϵ and ϵ' the angles which incident and refracted ray make with the surface normal.

Let us assume that we have the transformation equations given in the form

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = A\begin{pmatrix} x \\ y \end{pmatrix} + B\begin{pmatrix} n\xi \\ n\eta \end{pmatrix}, \qquad \begin{pmatrix} n'\xi' \\ n'\eta' \end{pmatrix} = C\begin{pmatrix} x \\ y \end{pmatrix} + D\begin{pmatrix} n\xi \\ n\eta \end{pmatrix} \tag{VIII, 1}$$

with

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} \bar{D} & -\bar{B} \\ -\bar{C} & \bar{A} \end{pmatrix} = \begin{pmatrix} \bar{D} & -\bar{B} \\ -\bar{C} & \bar{A} \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} E & 0 \\ 0 & E \end{pmatrix}. \tag{VIII, 2}$$

It is now our task to calculate A, B, C, D. What do we know that will help us? First, let us give another coordinate system which is connected with the surface and has its z^* axis in the direction of the normal. We have

$$\begin{aligned} x^* &= x^{*\prime}, \\ y^* &= y^{*\prime}, \end{aligned} \tag{VIII, 3}$$

as they are the coordinates of the incident point. Now this coordinate system can be obtained by rotating the previous one around the unchanged y (y') axis by the angles ϵ and ϵ', respectively. Thus, we get

$$\begin{aligned} x^* &= x\cos\epsilon, & n\xi^* &= n\xi/\cos\epsilon, \\ y^* &= y, & n\eta^* &= n\eta, \\ x^{*\prime} &= x'\cos\epsilon', & n'\xi^{*\prime} &= n'\xi'/\cos\epsilon', \\ y^{*\prime} &= y'\cos\epsilon', & n'\eta^{*\prime} &= n'\eta'/\cos\epsilon'. \end{aligned} \tag{VIII, 4}$$

What we need further is the connection between the directions of incident and refracted ray. This connection is given by the refraction law.

$$\begin{aligned} n'\xi^{*\prime} - n\xi^* &= (n'\cos\epsilon' - n\cos\epsilon)\xi_F, \\ n'\eta^{*\prime} - n\eta^* &= (n'\cos\epsilon' - n\cos\epsilon)\eta_F, \end{aligned} \tag{VIII, 5}$$

where ξ_F and η_F are the coordinates of the unit vector along the surface normal at the point x^*, y^*. From this we find easily, if ${}_Ir$ and ${}_{II}r$ are the main radii of curvature of the refracting surface, that

$$n\xi_F = \left(\frac{\cos^2\omega}{{}_Ir} + \frac{\sin^2\omega}{{}_{II}r}\right)x^* + \cos\omega\sin\omega\left(\frac{1}{{}_{II}r} - \frac{1}{{}_Ir}\right)y^*, \tag{VIII, 6}$$

$$n\eta_F = \cos\omega\sin\omega\left(\frac{1}{{}_{II}r} - \frac{1}{{}_Ir}\right)x^* + \left(\frac{\sin^2\omega}{{}_Ir} + \frac{\cos^2\omega}{{}_{II}r}\right)y^*.$$

If we eliminate the starred magnitudes, we get

$$\begin{aligned} x' &= \frac{\cos\epsilon}{\cos\epsilon'}x, \\ y' &= y, \\ n'\xi' &= \Gamma(\alpha x + \beta y) + \frac{\cos\epsilon'}{\cos\epsilon}n\xi, \\ n'\eta' &= \Gamma(\beta x + \gamma y) + \frac{\cos\epsilon}{\cos\epsilon'}n\eta. \end{aligned} \tag{VIII, 7}$$

Here we have set for abbreviation

$$\begin{aligned}\Gamma &= n' \cos \epsilon' - n \cos \epsilon,\\ \alpha &= \cos^2 \omega/{}_I r + \sin^2 \omega/{}_{II} r,\\ \beta &= \cos \omega \sin \omega(1/{}_{II} r - 1/{}_I r),\\ \gamma &= \sin^2 \omega/{}_I r + \cos^2 \omega/{}_{II} r.\end{aligned} \qquad \text{(VIII, 8)}$$

We see that this is, in general, the equation for a semitorted system, with

$$A = \begin{pmatrix} \dfrac{\cos \epsilon}{\cos \epsilon'} & 0 \\ 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

$$D = A^{-1} = \begin{pmatrix} \dfrac{\cos \epsilon'}{\cos \epsilon} & 0 \\ 0 & 1 \end{pmatrix}, \qquad C = \Gamma \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}. \qquad \text{(VIII, 9)}$$

These matrices fulfill the fundamental equations (II, 4), as the reader can easily verify. The system is orthogonal only if C is a diagonal matrix, which is the case if $\beta = 0$, or if

$$1/{}_I r = 1/{}_{II} r. \qquad \text{(VIII, 10)}$$

This signifies that the incident point is an umbilic point, or $\omega = 0$, which means that the incident plane coincides with one of the main sections of the surface. In all other cases we have a general semitorted system.

Now let us consider the transfer formulae. As we go from one surface to the next, we must shift the origin along the axis by the amount of the distance between the two incident points (d_{12}). Also, the meridian plane of the first surface does not necessarily coincide with that of the second surface. They might form an angle θ_{12}. We find that

$$\begin{aligned} x' &= \cos \theta_{12} x + \sin \theta_{12} y + (d_{12}/n)(\cos \theta_{12} n\xi - \sin \theta_{12} n\eta),\\ y' &= -\sin \theta_{12} x + \cos \theta_{12} y + (d_{12}/n)(\sin \theta_{12} n\xi + \cos \theta_{12} n\eta),\\ n'\xi' &= \cos \theta_{12} n\xi - \sin \theta_{12} n\eta,\\ n'\eta' &= \sin \theta_{12} n\xi + \cos \theta_{12} n\eta. \end{aligned} \qquad \text{(VIII, 11)}$$

The final transformation which gives A, B, C, D, for a system with a finite number of surfaces is the product of transformations similar to formulae (VIII, 7) and (VIII, 8). If we write instead of (VIII, 7)

$$\begin{pmatrix} x' \\ y' \\ n'\xi' \\ n'\eta' \end{pmatrix} = \Phi_1 \begin{pmatrix} x \\ y \\ n\xi \\ n\eta \end{pmatrix} \qquad \text{(VIII, 12)}$$

and instead of (VIII, 11)

$$\begin{pmatrix} x' \\ y' \\ n'\xi' \\ n'\eta' \end{pmatrix} = \Theta_{12} \begin{pmatrix} x \\ y \\ n\xi \\ n\eta \end{pmatrix}$$

we get for a system of κ surfaces

$$\begin{pmatrix} x' \\ y' \\ n'\xi' \\ n'\eta' \end{pmatrix} = \Theta_{\kappa,\,\kappa+1} \Phi_\kappa \Theta_{\kappa,\,\kappa-1} \Phi_{\kappa-1} \cdots \Theta_{12} \Phi_1 \Theta_0 \begin{pmatrix} x \\ y \\ n\xi \\ n\eta \end{pmatrix}.$$

Thus, in each special case we can easily calculate the matrices A, B, C, D.

The Effect of Pre-Exposure in Spectrum Photography

DANIEL NORMAN, *George Eastman Laboratory of Physics, Massachusetts Institute of Technology, Cambridge, Mass.*

(Received September 14, 1936)

The effect of pre-exposure on Eastman 40, 50, I–G, I–F, and Cramer Hi-Speed plates was investigated and was found to be similar on all five. The optimum pre-exposure density in every case, as was to be expected, was equal to the density at the beginning of the straight line portion of the *D*-log *I* curve. To eliminate the Herschel effect, it was found necessary to pre-expose with a light source having an effective wave-length longer than the wave-length of any portion of the spectrum which was to be used for density measurements. A definite fading of the latent image amounting to a density decrease of 0.07 at a density level of 0.8 was observed in plates which were developed 12 hours after exposure. This fading ceased, in the cases studied, after about twenty hours.

THE value of a pre-exposure in the photography of faint spectra was known as far back as 1908, when Wood[1] found that it almost doubled the sensitivity of the emulsions he was using, and that the most effective pre-exposure density was that which equalled the density at the beginning of the straight-line portion of the *D*-log *E* curve. At that time he wrote: "Having used it (pre-exposure) with very great success I think it is worth bringing anew to the attention of those engaged upon the work of photographing weak spectra," but for many years very little use was made of the effect.

In 1927 Weinland[2] showed that the chief intermittency failure lies in the underexposed region of the *D*-log *I* curve and pointed out that pre-exposure of the plate practically eliminates the effect. Since then pre-exposure has been used in stellar spectroscopy[3] both to reduce the intermittency effect and to increase the intensity of the underexposed regions of the spectrum so that they become measurable by the microphotometer. Spectroscopists in general, however, have apparently made little use of the effect, although it should be of great value in photographing Raman and other spectra where the available light intensities are low. It may be remarked here that pre-exposure does not actually change the sensitivity of the silver halide grains but merely increases the density of the underexposed portions of the spectrum relative to the fully exposed portions.

In the present study of the effect of pre-exposure on such commonly used emulsions as Eastman 40, 50, I–G, I–F, and Cramer Hi-Speed Special plates, the result of changing the wave-length of the light source used in pre-exposing was first investigated. It was found that when the wave-length of this light was shorter than the wave-length of a line whose density was to be measured, an actual decrease in density occurred at density levels of 1.3 to 0.9—presumably due to the Herschel effect. If a plate is to be pre-exposed and still used for photometric purposes, it is necessary that the pre-exposing light be of wave-length greater than that of any part of the spectrum which is to be photometered. An E.K. Co. Cine-Red filter was therefore used before the pre-exposing light source, since it transmits a negligible amount of light below 5900A, and no lines of wave-length greater than 5460A were measured.

The prefogging apparatus consisted of a 6-watt, 110-volt lamp, behind two thicknesses of glass ground on both sides and the Eastman Cine-Red filter. This light source was placed two meters from a special plateholder which permitted the exposure of a 2×10 inch section of a 4×10 inch plate. Both light source and plateholder were contained in a single light-tight box. No irregularity could be detected in the density of the fog produced by this apparatus. The exposure times ranged from 1 second to 140 seconds.

The plates were calibrated in a Hilger E″ quartz prism spectrograph, using calibrated neutral wire screens[4] to control the light intensity. The light source was a quartz mercury arc run at constant temperature and amperage, permanently fastened two meters from the

[1] R. W. Wood, Astrophys. J. **27**, 379 (1908).
[2] C. E. Weinland, J. O. S. A. **15**, 337 (1927).
[3] See, for example, F. L. Whipple, *Lick Observatory Bulletin No. 442.*
[4] G. R. Harrison, J. O. S. A. **18**, 492 (1929).

spectrograph slit. No lens was used, and the mechanically oscillated wire screens were placed immediately before the slit. An electromagnetic shutter gave 17 second exposures which were identical (to better than 0.01 second).

Twenty spectra were impressed on each plate by means of ten different screen combinations, each screen combination being used once on the prefogged and once on the unfogged portion of the plate. To avoid any spurious effects which might arise from the gradual gradation in intensity of the successive spectra, the order of screen combinations was staggered so that the intensity sequence was 8.9, 40, 3.7, 69, 1.8, 14, 100, 21, 60, and 26 percent.

All plates were developed six minutes at 18°C in the Eastman D11 formula. Fresh developer was used for each plate, and the plates were brushed during development to reduce the Eberhard effect.

The photovoltaic densitometer used to measure the plates was linear for densities less than two, and all density measurements could be duplicated to 0.01.

Measurements were made at six different wave-lengths, from 2400A to 5460A, on each plate and several sequences of plates were taken in which all the variables but one were held constant. The factors varied were the emulsion, the density of the pre-exposure, and the interval between the pre-exposure and the impressing of the standards.

As is commonly known,[5] the effect of pre-exposure is to increase the sensitivity but to decrease the contrast of an emulsion (plotting D-log I curves). The five emulsions used all behaved in exactly the same way, and the conclusions reached apply to all of them.

A pre-exposure density of 0.05 caused an average density increase of 10 percent at all densities less than 1.7, at all wave-lengths (on Eastman 50's). Both the percentage and the numerical increases were largest at low densities and decreased progressively at higher densities.

The pre-exposure density was varied from 0 to 1.2, but at the higher densities the contrast was so low that the plates could not be measured in a densitometer. With a contrast microphotometer[6] it might be possible to utilize high pre-exposure densities. For ordinary work the optimum pre-exposure density is equal to the density at the beginning of the straight-line portion of the D-log I curve. Since this is true for all five emulsions at all wave-lengths, and since it is true for the (unnamed) emulsion used by Wood[1] it is quite probable that this is a rule which is applicable to all emulsions.

In no case was the increase of density of lines in the prefogged portion of the plate over the corresponding lines in the unfogged portion of the plate greater than the pre-exposure density. In the cases where the pre-exposure density was too low to be measured, its effect was negligible.

Since it is not always possible to give a controlled pre-exposure immediately before photographing a spectrum, it seemed of interest to investigate the effect of increasing the interval between the pre-exposure and the impressing of the standards on the plate.

The interval between the pre-exposure and the regular exposure was varied from three to thirty-eight hours. A decrease in density was observed which amounted, in Eastman 50's, to 0.07 in thirteen hours (at a density level of 0.8) compared to plates in which the interval was three hours. The amount of fading became vanishingly small after twenty hours.

The amount of fading was different at different wave-lengths, but was nearly constant at different densities of the same wave-length. It was also constant at any wave-length for different plates of the same emulsion for the same interval. The increase of density of the latent image reported by Bullock[7] was not observed, presumably because it occurred within a relatively short time after exposure, and no intervals less than three hours were obtained in this investigation.

The author wishes to thank Professor G. R. Harrison for the use of the facilities of his laboratory. He is also greatly indebted to Dr. Fred L. Whipple of the Harvard Observatory for advice.

[5] See, for example, Hardy and Perrin, *The Principles of Optics* (McGraw-Hill, 1932), p. 215.

[6] F. W. Sears, J. O. S. A. **25**, 162 (1935).

[7] E. R. Bullock, Abridged Scientific Publications from the Kodak Research Laboratories **14**, 173 (1930); **15**, 45 (1931).

Changes in Color Temperature of Tungsten-Filament Lamps at Constant Voltage*

DEANE B. JUDD, *National Bureau of Standards, Washington, D. C.*

(Received September 23, 1936)

Seasoned lamps, operated at constant voltage, gradually decrease in color temperature because of evaporation of the filament. Since the current does not decrease enough to correspond to an appreciable change in filament temperature, this decrease in color temperature is ascribed in large part to the accumulation of the familiar brown film on the inside surface of the bulb. The size and shape of the bulb is, therefore, important; each type must be separately investigated. Results are given for five types. The effect of operating seasoned lamps is a gradual decrease in color temperature, linear with time, which, like the rate of evaporation of tungsten, was found for 400-watt projection lamps, to be nearly proportional to the thirtieth power of the temperature. The effect of seasoning new lamps for one hour at about rated voltage was a rise in color temperature from 3050° to about 3140°K, approximately half of which occurred in the first three minutes.

CONTENTS

I. Introduction 409
II. Theoretical considerations 409
A. Filaments of uniform diameter 410
B. Filaments of nonuniform diameter 411
III. Apparatus 413
A. Devices for operating lamps at constant voltage 413
B. Measurement of luminous flux 413
C. Measurement of color temperature 413
IV. Method 413
V. Reduction and summary of data 414
VI. Discussion and supplementary results 416
A. Dependence of lamp permanence on filament temperature 416
B. Type of lamp for color-temperature standards 417
C. Alternating-current *versus* direct-current operation 417
D. Correlation between luminous efficiency and color temperature of individual lamps in the life test 418
E. Effects of seasoning the lamps 418
F. Supplementary measurements of the brown deposit 419
VII. Conclusions 420
VIII. Acknowledgments 420

I. INTRODUCTION

WHEN a seasoned incandescent lamp is burned at constant voltage, its chromaticity is not quite constant but changes slowly toward yellow. This chromaticity is conveniently and customarily specified by color temperature, defined as the temperature which an inclosure would need have in order to yield the same chromaticity by thermal radiation. The light output of the lamp also gradually decreases.

This investigation was begun in 1930 with the primary purpose of discovering which of a number of types of tungsten-filament incandescent lamps was best suited for standards of color temperature. It was also expected that the study would show how much a working standard could be used before a check against a primary standard was advisable. Furthermore, the suggestion was made that measurements of luminous efficiency might serve as a check on the color-temperature constancy of lamps, and therefore an attempt has been made to correlate the decline in luminous efficiency with the decline in color temperature.

The types of lamp investigated are those which were thought of as possible best choices for color-temperature standards; Table I gives a description of these types, the number tested, the type of current used, the initial color temperature, and the date and duration of the test.

The general plan of the investigation has been to burn the lamps at constant voltage and to measure at intervals, (1) current, (2) light output, and (3) color temperature.

It is the purpose of this paper to describe the apparatus and method and to give the essential results with some discussion of their significance.

II. THEORETICAL CONSIDERATIONS

The change in color temperature of a tungsten-filament lamp at constant voltage is ascribable to evaporation of tungsten from the filament. Such evaporation may change the color temper-

* Publication approved by the Director of the National Bureau of Standards of the U. S. Department of Commerce.

TABLE I. *The types of lamp investigated, their number, duration of test, and color temperature.*

Rated Watts	Type of Bulb	Type of Filament	Number Tested	Duration of Test, Hours	Type of Current	Initial Color Temperature	Test Started
400	Tubular, gas-filled	Monoplane	5	64	A.C.	3100°K	6/18/32
			5	64	D.C.	3100	6/18/32
			5	700	A.C.	2850	8/ 3/32
			5	700	A.C.	2360	8/ 3/32
500	Pear-shaped, gas-filled	Crown (C7)	5	128	D.C.	2850	5/ 4/31
500	Pear-shaped, gas-filled	Circular-coil (C7A)	5	128	D.C.	2840	10/19/31
60	Pear-shaped, vacuum	Squirrel-cage	5	128	D.C.	2360	5/ 4/31
60	Pear-shaped, vacuum	Circular-coil (C9)	5	128	D.C.	2360	10/19/31

ature of the lamp in two ways. It may cause a change in the true temperature of the filament, itself, and it may change the color temperature of the lamp by a deposit on the inside of the bulb acting as a selective filter.

The properties of tungsten and tungsten filaments have been thoroughly investigated and summarized by Forsythe and Worthing[1] and by Jones and Langmuir.[2] From these investigations the change in color temperature of the filament, itself, from the loss of tungsten may be closely estimated regardless of whether or not the diameter remains uniform.

A. Filaments of uniform diameter

Consider, as is usually done, an ideal straight tungsten filament of circular cross section having a smooth, polished surface. Let there be no cooling at the leads nor any cooling by conduction to gases in contact with the filament. For such a filament of constant length L, with a constant voltage V applied over this length, it has been shown by Jones and Langmuir that temperature of the filament T and current I flowing in it are proportional to certain powers of the diameter d, so that we may write:

$$T/T_0=(d/d_0)^n, \tag{1}$$

where n varies from 0.168 to 0.176 within the temperature range 2300 to 3000°K, and

$$I/I_0=(d/d_0)^m, \tag{2}$$

where m varies from 1.80 to 1.79 within the same range. The subscript 0 denotes the values at the beginning of the burning test.

The color temperature θ differs from the true temperature of the filament by only 50 to 100 degrees throughout this range of temperature,[1] so without appreciable error for our purpose Eq. (1) may be written:

$$\theta/\theta_0=(d/d_0)^{0.17}. \tag{3}$$

Assume that the diameter of the filament is uniform and decreases uniformly because of evaporation from the surface and linearly with time t in service, thus:

$$d/d_0=1-t/t' \tag{4}$$

where t' is the time required for the complete evaporation of the filament.

For a lamp whose change in color temperature is ascribable wholly to change in true temperature of a uniform filament it could be expected from Eqs. (3), (2) and (4) that the color temperature and current would vary with time as follows:

$$\theta/\theta_0=(1-t/t')^{0.17}, \tag{5}$$

$$I/I_0=(1-t/t')^{1.8}, \tag{6}$$

that is, the color temperature should decrease approximately as the fifth root of time and the current should decrease approximately as the square.

Actual filaments, however, are not uniform to start with, and it may be readily seen that a

[1] W. E. Forsythe and A. G. Worthing, "The Properties of Tungsten and the Characteristics of Tungsten Lamps," Astrophys. J. **61**, 146 (1925).

[2] H. A. Jones, and Irving Langmuir, "The Characteristics of Tungsten Filaments as Functions of Temperature," Gen. Elec. Rev. **30**, 310–319, 354–361, 408–412 (1927).

narrow, constricted region of the filament will cause development of a "hot spot" which will be always at a higher temperature than other parts of the filament and which, barring other accidents, will ultimately cause its failure.[3] Irregularities in the filament also appear because of formation of large crystals.[4] The filaments of the lamps burned for 700 hours show considerable evidence of such crystallization.

B. Filaments of nonuniform diameter

Consider changes in the original filament such that a certain fraction f of the length is reduced in diameter to a certain fraction c of the original diameter. To trace the effect of these changes in dimensions requires three separate sets of symbols as follows:

(1) Voltage V, applied to the original filament, diameter d_0 produces current I_0, color temperature θ_0 and resistivity ρ_0.

(2) Voltage V, applied to the diminished filament produces in the unreduced part, diameter $d'=d_0$, a lower current I', a lower color temperature θ' and a lower resistivity ρ'.

(3) And it produces in the reduced part, diameter $d=cd_0$, the same current $I''=I'$ and, as will appear presently a higher color temperature θ'', and a higher resistivity ρ''.

The resistance R of the partially reduced filament may be expressed in terms of the resistivities of the two component parts, thus:

$$R=L/d_0^2[\rho'(1-f)+\rho''(f/c^2)]. \quad (7)$$

The current flowing through the partially reduced filament is: $I=I'=V/R$; that flowing through the original filament is: $I_0=Vd_0^2/L\rho_0$; and the ratio from Eq. (7) may be written

$$I/I_0=I'/I_0=\rho_0/[\rho'(1-f)+\rho f/c^2]. \quad (8)$$

Assuming that no heat is transferred from the narrow part of the filament to the unreduced part by conduction, we may write from Langmuir and Jones' table of relations between the variables of ideal tungsten filaments[2] for constant current since $I=I'$:

$$\theta/\theta'=(d_0/d)^{0.9}=1/c^{0.9}; \quad (9)$$

also

$$\rho/\rho'=1/c^{1.1}. \quad (10)$$

The effect of introducing a narrow place in the filament is, of course, to decrease the current, but as shown below even with this decreased current (except for f nearly unity) there is a rise in the color temperature of the narrow part of the filament.

The color temperature of the filament as a whole is a mean of the color temperatures of the two component parts. Use has been made of the red-blue ratio[5] taking wave-lengths 660 and 470 mμ, for estimating the mean. The total amount of radiant energy at 660 mμ from both the narrow and unchanged parts of the filament was computed by Wien's law from the respective color temperatures, a similar total at 470 mμ found, and the ratio computed. The color temperature of the filament as a whole was taken as the temperature of the Wien radiator having this same red-blue ratio.

As an illustration of the dependence of color temperature on the degree of irregularity with which the filament decreases in diameter, there have been computed the color temperatures of five filaments each differing from the original filament by losing enough tungsten to reduce the current with constant voltage V by one percent. The first of these five filaments has its diameter reduced uniformly over the entire length ($f=1$), the second is unchanged over nine-tenths of its length, but has its diameter uniformly decreased over the remaining one-tenth ($f=0.1$), and so on for $f=0.05$, 0.02 and 0.01. The details of the computation and the results are shown in Table II.

The reduction in diameter c required to reduce the current from I_0 to $0.99I_0$ was found from Eqs. (8) and (10):

$$c^{3.1}=\rho' f/\rho_0/[0.99-\rho'(1-f)]. \quad (11)$$

The resistivities ρ' and ρ_0 refer to filaments of the same diameter ($d'=d_0$) but the filament of resistivity ρ' carries one percent less current. From relations given by Jones and Langmuir, the temperatures, and hence the color temperatures, of the filaments, are closely proportional to the 0.6 power of the current for diameter constant:

$$\theta'/\theta_0=(I'/I_0)^{0.60} \quad (12)$$

[3] R. Becker, "Lebensdauer und Wolframverdampfung," Zeits. f. tech. Physik **6**, 309 (1925).

[4] G. R. Fonda, "Burn-Out of Incandescent Lamps," Gen. Elec. Rev. **32**, 206 (1929).

[5] W. E. Forsythe, "Color Match and Spectral Distribution," J. O. S. A. and R. S. I. **7**, 1115 (1923).

TABLE II. *Change in color temperature according to distribution of tungsten removed from filament.* (These computations refer to a one percent decrease in current, $I''/I_0=0.99$.)

Fraction of Original Length, L, Reduced in Diameter f	Fraction of Original Diameter Required Over Fractional Length, f, to Produce a 1% Decrease in Current see Eq. (14) c	Fractional Increase in Color Temperature at Narrow Part of Filament $\theta''/\theta_0=0.994^*/c^{0.9}$ θ''/θ_0	Color Temperature of Unchanged Part of Filament, θ', of Narrow Part, θ'', and of Whole Filament, $\theta_{R/B}$ for Two Initial Color Temperatures, θ_0, in Degrees Kelvin					
			For $\theta_0=2400$°K			For $\theta_0=2800$°K		
			θ''	θ' See Note†	$\theta_{R/B}$	θ''	θ' See Note†	$\theta_{R/B}$
1.00	0.995	0.999	2398	—	2398	2797	—	2797
.10	.950	1.042	2501	2386	2403	2918	2783	2801
.05	.909	1.084	2602	2386	2407	3035	2783	2804
.02	.818	1.191	2858	2386	2422	3335	2783	2814
.01	.724	1.330	3192	2386	2460	3724	2783	2840

* The factor, 0.994 is introduced because Eq. (3) requires θ''/θ_0 to equal 0.999 for $c=0.995$; this is a drop in color temperature with voltage and length constant due to decreased current; it does not follow from Eq. (9) which refers to constant current.

† The length of filament, $(1-f)L$, unchanged in diameter has a color temperature lower than θ_0 because of the one percent decrease in current, thus: from Eq. (12): $\theta'/\theta_0=(I'/I_0)^{0.60}=0.99^{0.60}=0.994$.

and resistivity is closely proportional to the 1.2 power of the color temperature regardless of filament dimensions:

$$\rho'/\rho_0=(\theta'/\theta_0)^{1.2}. \tag{13}$$

These exponents vary slowly with color temperature much as those of Eqs. (1) and (2); but, as before, values representative of 2800°K are used for simplicity, the errors resulting from neglect of the more accurate relations being of no consequence for the present purpose.

From Eqs. (12) and (13) may be written:

$$\rho'/\rho_0=(I'/I_0)^{0.72}$$

which for I'/I_0 equal to 0.99 is approximately 0.993. Substitution of this value in Eq. (11) gives the relation from which the values of c in the second column of Table II were found:

$$c^{3.1}=0.983f/(0.017+0.983f). \tag{14}$$

It may be noted that the numerical results computed from Eq. (14) differ only slightly from those obtained from the simpler relation:

$$c^2=f/(0.01+f) \tag{14a}$$

found from Eq. (8) by neglecting the differences between ρ_0, ρ', and ρ''.

The third column gives the ratio of color temperature of the "hot spot" to the original color temperature, θ/θ_0, according to a relation found from Eqs. (3) and (9). There are also given for two initial color temperatures, 2400°K and 2800°K, the color temperature of the "hot spot" f, the color temperature of the remainder, $1-f$, of the filament, and the color temperature of the whole filament, $\theta_{R/B}$, estimated by the red-blue-ratio method.

It may be seen for $f=1$ that a reduction in diameter of the whole filament by one-half of one percent is required to produce a one percent reduction in current. This reduction may be expected to lower the color temperature by but 0.1 of one percent (2 or 3 degrees); this is in accord with Eq. (3). The remainder of the entries in the table show that the color temperature of the whole filament is relatively independent of how the reduction in diameter required to reduce the current by one percent is distributed, the effect of deviating from uniform reduction in diameter ($f=1$) being to raise the color temperature slightly. Only when the whole decrease is limited to one-twentieth or less of the filament length does the increase amount to more than 9°K. Such an irregularity produces a rather severe "hot spot," a locality in which the color temperature rises about 200°K above that of the remainder of the filament. The extreme restriction, $f=0.01$, was included as representative of a lamp burning out from development of a "hot spot."

These results apply to ideal straight filaments in a vacuum; most of the lamps actually tested

had coiled filaments in gas-filled bulbs. It is believed, however, that the considerations apply essentially to such lamps also. We may expect therefore that the color temperature of the filament will in general rise slowly with continued operation, and even in the unlikely case of a very uniform filament the fractional decrease in color temperature is expected not to be any more rapid than the tenth root of the fractional decrease in current (see Eqs. (5) and (6)). This permits a considerably simplified discussion of the results which may conveniently be mentioned now. As will appear later, decreases have been found in color temperature of as much as 35°K accompanied by decreases in current up to one percent. From Table II it may be seen that the maximum decrease ascribable to change in color temperature of the filament is 3° at 2800°K. It may be concluded, therefore, that by far the largest part of the decrease in color temperature found is ascribable to the accumulation of the brown film on the inside of the bulb. No further attention will be paid to possible changes in color temperature of the filament because there is no evidence that such changes have significantly influenced the experimental results.

This conclusion emphasizes, however, that size, shape and gaseous content of the bulb determine the permanence of the lamp as a color-temperature standard. A large bulb provides a larger glass surface on which the film may be deposited with a consequent reduction in its thickness. By variation in shape of the bulb convection currents within may be controlled so that most of the film is deposited on parts of the bulb off the line of sight. Each type of lamp must therefore be separately investigated.

III. Apparatus

The apparatus consisted of devices for burning lamps at constant voltage, and devices for measuring luminous flux, and color temperature at known voltages.

A. Devices for operating lamps at constant voltage

For most of the work the lamps were operated on alternating current at prescribed voltages in the equipment developed for life testing of incandescent lamps. Since lamps used for standards of color temperature are customarily operated by direct current, the early tests were made with direct current until a comparison of the two methods showed little significant difference. In each case the voltage was controlled to within 0.25 volt.

B. Measurement of luminous flux

The light output of each lamp was either measured by means of an 88-inch photometric integrating sphere and visual photometric equipment used in routine tests at the Bureau for many years, or by a 60-inch sphere equipped with a caesium photoelectric cell in a Sharp and Smith circuit.[6] It has been found possible in routine measurements to detect variations in luminous flux as small as 0.25 of one percent with the photoelectric photometer which is a precision considerably higher than was obtained with the visual photometer.

C. Measurement of color temperature

The color temperature of the lamps was measured by comparison with lamp standards on the Nela scale of color temperature[7] according to the substitution method on the color comparator described in 1930[8] either with the flat 16° circular photometric field divided vertically in the center and viewed binocularly, or with the Lummer-Brodhun contrast photometer head with the contrast strips removed giving a $9\times13°$ field with a complex double-trapezoid pattern viewed monocularly. The latter more complex field gave no higher precision but was preferred because it largely eliminated errors due to changes in chromatic sensitivity of one-half of the observer's retina relative to the other half over time intervals of the order of ten minutes.

IV. Method

The method for seasoned lamps was (1) to set them for the desired initial color temperature,

[6] C. H. Sharp and H. A. Smith, "Further Developments in Photoelectric Photometers," Trans. I. E. S. **23**, 434 (1928).

[7] This work was done in 1930 to 1932 prior to the establishment of the present scale of color temperature (see reference 12). The Nela scale differs so little from the present scale that no useful purpose would be served in transferring these results.

[8] D. B. Judd, "Precision of Color Temperature Measurements Under Various Observing Conditions; A New Color Comparator for Incandescent Lamps," Nat. Bur. Stand. J. Research **5**, 1161 (1930); RP252.

θ_0 (see Table II) by adjustment to voltage V, (2) to check for possible seasoning effects by measurement of current and color temperature at this voltage immediately after this period (about three minutes) of adjustment, (3) to operate them at this voltage for the following time intervals in hours, 1, 2, 4, 8, 16, 32, 64, 50, 50, 50 . . ., continuing until the durations given in Table I had been reached and (4) after each such time interval to measure current, luminous flux and color temperature.

The method for unseasoned lamps was the same except that they were seasoned for one hour at 3140°K and the check for seasoning effect was applied once during this hour after about seven minutes, the initial adjustment requiring longer than three minutes because of the rapid rise in color temperature during initial stages of seasoning. For two lamps the seasoning effect was studied continuously for 10 minutes by settings of color temperature every 30 seconds.

V. Reduction and Summary of Data

The 400-watt lamps were taken from stock; the 60-watt, squirrel-cage filament lamps were taken from lamp sockets after actual use of unknown amount; the remainder of the lamps (see Table I) were purchased as photometric standards. As might be expected, only the lamps taken from stock showed any evidence of a seasoning effect.

Fig. 1 shows typical variation of color temperature with time in service for unseasoned lamps. The seasoning effect found is a rapid rise of 50 to 100°K in color temperature accompanied in every case, contrary to expectation, by a one percent decrease in current. The color-temperature rise is best shown on Fig. 1 by the plot with the enlarged time scale. The effect of burning seasoned lamps is a slow decline in color temperature with continued but slow decrease in current. The plot to the intermediate time scale shows that seasoning is completed some time in the first hour of burning. This result justified the usual seasoning procedure which is to operate such lamps at about this color temperature (3140°K) for an hour. The color temperature rises during the first minute of seasoning by an amount about equal to the decrease produced by 50 hours of burning after seasoning; hence color temperature during the initial stages of the seasoning process changes several thousand times as fast as it does during operation of a seasoned lamp. This rapid rise in color temperature is customarily ascribed to change in surface condition of the filament so as to increase within the visible spectrum the spectral emissivity of short wave radiant energy relative to that for long wave.

Lamps other than those taken from stock showed only the gradual decrease in color temperature characteristic of seasoned lamps. It is important to note that the color-temperature time curve closely approximates a straight line. Analysis of the data shows that the actual deviations from a straight line are not significant. For the mean of these ten lamps, the deviations from a straight line are in each case less than 3 degrees; they are ascribable to experimental error, apparently in large part to error in control by the lamp standard of color temperature. In addition to this, the individual lamps whose mean is represented on Fig. 1 gave a spread of about 20°, but this is due chiefly to uncertainty in the initial determination of voltage for 3100°K because of the rapidity of the seasoning effect during the first five minutes. The curves for the individual lamps showed generally parallel courses; significant deviations could be traced to the fusing of two adjacent coils of the filament accompanied by an increase in current, to development of an unusually heavy region of brown deposit in the line of sight, or similar incidents. Significant erratic variations not accompanied by these incidents, such as reported by Howe[9] and Dziobek,[10] were not found with these lamps. The previously noted erratic variations would seem from the present tests to be ascribable to undetected sources of error; but it is possible that the present results fail to be typical because of insufficient number of lamps tested.

The other types of lamps and the same type at lower initial color temperatures likewise

[9] H. E. Howe, "The Color Temperature of the Gas-Filled Tungsten Lamp as a Function of Time in Service," J. O. S. A. and R. S. I. **13**, 304 (1926).

[10] W. Dziobek, "Über die Verwendbarkeit der Wolfram-Vakuumlampe zu sensitometrischen Messungen," Zeits. f. Wiss. Photo. **31**, 96 (1932).

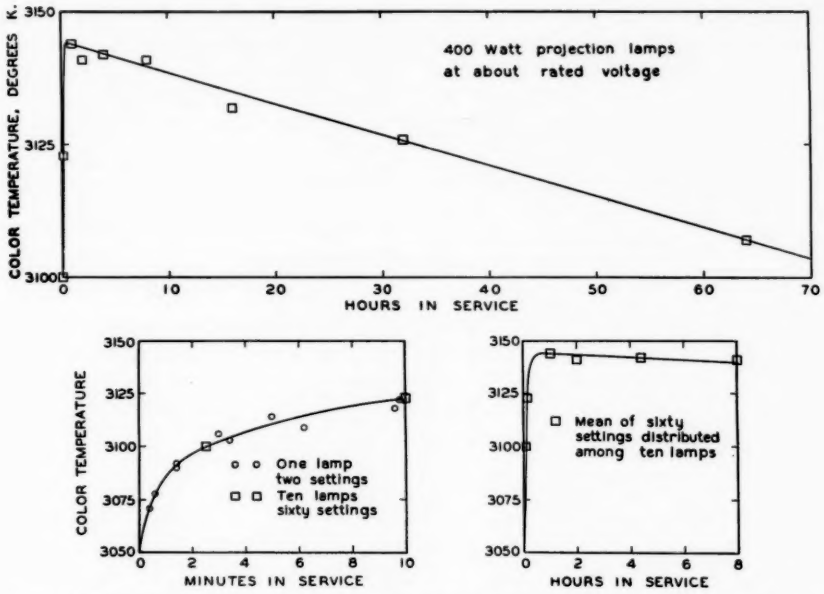


FIG. 1. Color temperature of unseasoned lamps as a function of time in service. The color temperature rises rapidly during seasoning, then declines slowly.

exhibited a decrease in color temperature not significantly nonlinear with time, though, of course, the rapidity of color-temperature change at the lower temperatures was notably less than at 3140°K. It is concluded, therefore, that the permanence of a lamp for use as a standard of color temperature at a given voltage can be satisfactorily indicated by the average decline in color temperature per unit time.

Luminous efficiency and current of the lamps studied were also found to decrease with time in a nearly linear way; this agrees with results found for gas-filled lamps used as photometric standards.[11]

The method used in reducing the data was accordingly to compute the change per hour by reference to the average current, luminous efficiency, and color temperature for the group of lamps, no further attention being paid to values for individual lamps. In the case of current the method was merely to divide the difference between the initial and final values by the number of hours the lamps had been burned; this gives a result of the right order of magnitude. But, with luminous efficiency and color temperature least-square solutions were found for the constants a and b in the expression, $a+bt$; the values of b thus found expressed as the change per hour are given in Table III, together with uncertainties of these values. The uncertainties are taken as 4.9 times the probable error obtained in the customary way from the departures of the actual data from the linear relation found by least squares. On the assumption that the lamps studied are fair samples of all lamps of the same type, these uncertainties may be interpreted as limits which would be exceeded on the average only once in a thousand such tests as these.

Table III is a summary of these results including the uncertainties in rates of change of luminous efficiency and color temperature with duration of the test. The data cannot be correctly interpreted without a consideration of these uncertainties which in some cases are undesirably large. Note, for example, that the projection lamps were found at 2850°K to decline in color temperature at the rate of from 0.05 to 0.07 degrees per hour, while the new standard type at about the same temperature yielded from 0.02 to 0.14 degree. The conclusion is therefore that

[11] R. P. Teele, "Gas-Filled Lamps as Photometric Standards," Trans. I. E. S. **25**, 78 (1930).

there is no significant difference in permanence between these two types of lamps; this is borne out by a consideration of the corresponding decline in luminous efficiency. In another example, however, the projection lamps at 2360°K showed an average decline greater than 0.001 and less than 0.011 degree per hour, while the standard vacuum type lamps yielded an average between 0.01 and 0.09. The conclusion is that the superior permanence of the projection lamps at 2360°K to vacuum lamps at the same color temperature has been established to a high degree of probability; consideration of the corresponding results for luminous efficiency establishes this superiority beyond reasonable doubt because in this case there is no overlapping whatsoever.

In general, the uncertainty in rate of decline is greater for color temperature than for luminous efficiency. This is ascribable partly to the relatively less dependence of total luminous flux on distribution of the brown deposit on the bulb, but it is chiefly ascribable to the generally higher precision of the measurements of luminous flux. In the first part of the work, changes in the lamps could be detected much better (perhaps five times as well) by following the luminous flux than by following the color temperature. After two years of practice in technique of setting for color-temperature match and after several refinements in method and equipment including installation of the Lummer-Brodhun contrast photometer head, it was found possible to do about equally well with either method. At the end of the work measurements of luminous flux were again found somewhat superior for tracing small changes in the lamps, the increased precision being ascribable chiefly to the substitution of the photoelectric cell for the visual photometer.

VI. Discussion and Supplementary Results

A. Dependence of lamp permanence on filament temperature

It should first be pointed out that the rate of decline in both color temperature and luminous efficiency indicated in Table III varies approximately according to the thirtieth power of the temperature of the filament. The rate of vaporization of tungsten[2] has a similar dependence. The results are therefore in approximate accord with the view that the decline of both color temperature and luminous efficiency of these lamps is ascribable to formation of a brown film of tungsten or some tungsten compound on the inside of the bulb. Table IV shows a comparison of the observed values of these rates with powers of the true temperature of the filament; it indicates that within the uncertainty of the results color temperature of the lamp declines according to the twenty-third power of the temperature of the filament, and that luminous efficiency declines closely according to the thirty-first power of the temperature; but the agreement is not perfect. The indication is that even the small changes observed at color temperature 2360°K are somewhat larger than would be expected from the rate of vaporization of tung-

Table III. *Current, luminous efficiency and color temperature for the types of lamp studied, and their time rates of change when the lamps are operated at constant voltage V; see Table I for further description of lamps.*

Identification of Lamp Type	Voltage V	Current I_0 Amperes	dI/dt Milli-amperes per Hour	Luminous Efficiency ϵ_0 Lumens per Watt	$d\epsilon/dt$ Lumens per Watt per Hour	Color Temperature θ_0 Degrees Kelvin	$d\theta/dt$ Degrees Kelvin per Hour
Projection	110	3.82	−0.6	24.6	−0.058 ±0.007	3140	−0.6 ±0.1
	85	3.31	−0.03	16.0	−0.0026 ± .0004	2850	−0.06 ± .01
	52	2.54	−0.003	5.7	−0.00015 ± .00011	2360	−0.006 ± .005
Old Standard Type (C7)	110	4.20	+0.04	16.5	−0.0029 ± .0011	2850	−0.09 ± .05
New Standard Type (C7A)	100	4.08	+0.02	16.1	−0.0026 ± .0011	2840	−0.08 ± .06
Commercial Vacuum Type	105	0.475	−0.01	7.4	−0.0018 ± .0009	2360	−0.01 ± .08
Standard Vacuum Type	95	0.469	0.00	7.4	−0.0015 ± .0005	2360	−0.05 ± .04

TABLE IV. *Dependence on temperature of filament of rate of decline of luminous efficiency, ϵ, and color temperature, θ, of the lamp.*

Color Temperature θ	True Temperature T^*	$\frac{-0.59T^{31}}{10^{109}}$	$\frac{d\epsilon}{dt}$	$\frac{-0.447T^{23}}{10^{80}}$	$\frac{d\theta}{dt}$
3140	3046	−0.058	−0.058 ±0.006	−0.59	−0.6 ±0.1
2850	2768	−0.0030	−0.0026 ± .0004	−0.065	−0.06 ± .01
2360	2310	−0.00003	−0.00015± .00011	−0.0010	−0.006± .005

* Taken from data by Forsythe and Worthing (see reference 1) connecting true temperature with color temperature on the Nela scale.

sten. It is evident that other factors not taken into account have a significant influence on the results. Evaporation of material from the leads and filament supports may produce this effect; also the higher pressure of the gas may interfere with evaporation of tungsten from the filament at the higher temperatures.

The fact that the color-temperature change varies more slowly with temperature (twenty-third power) than rate of vaporization (thirtieth power and higher) suggests that at the high temperatures the convection currents of gas within the bulb are such as to favor, relative to low temperatures, the deposition of tungsten off the line of sight. This explanation does not apply to luminous efficiency because the light output refers to all directions from the filament rather than only one, nor is this explanation needed because luminous efficiency declines much more closely (thirty-first power) according to rate of vaporization of tungsten.

B. Type of lamp for color-temperature standards

From Table III it may be noted that the gas-filled lamps specially constructed for photometric standards were not found to be any more permanent in color temperature near 2850°K than the 400-watt projection lamps. Indeed, the projection lamps seem to be slightly more permanent although the difference is not certain. The reverse had been expected because the projection lamps had considerably smaller bulbs which should cause a thicker brown deposit in the line of sight if the distribution of deposit were uniform. The greater permanence of the projection lamps is now ascribed to the tubular shape of the bulbs which seems to cause a greater proportion of the deposit to occur out of the line of sight.

It may also be noted from Table III that, as is to be expected, the projection lamps are definitely more permanent at 2360°K than the vacuum lamps tested at the same temperature.

The projection type of lamp was therefore chosen for new standards of color temperature over the whole scale.[12] Advantages gained by this choice apart from the gain in permanence just noted are: (1) a single lamp can be used over the whole scale, (2) the projection lamps are more convenient because they are smaller and burn base down, and (3) they are cheap because they are produced in large quantities for another purpose.

C. Alternating-current *versus* direct-current operation

Five of the ten lamps burned at about 3100°K were operated on direct current, the other five on alternating current. When the results on these two groups of lamps were plotted it was decided from visual inspection of the graphs that there was no significant difference between the two kinds of operation. Alternating current did not, it is true, produce quite as much lowering of color temperature as direct current. For three reasons, however, this result at the time seemed incredible: (1) it had been suspected that a small effect in the reverse sense might show up because of possible fluctuations (60 per second) in temperature from alternating-current operation, the vaporizing action of which might therefore exceed that of direct-current operation at the same color temperature, (2) no difference in the decline of luminous efficiency was noted between the two groups, and (3) the difference observed seemed ascribable to chance. Hence in subsequent burning the less convenient method using direct

[12] H. T. Wensel, D. B. Judd and Wm. F. Roeser, "Establishment of a Scale of Color Temperature," Nat. Bur. Stand. J. Research **12**, 527 (1934); RP677.

TABLE V. *Alternating-current operation compared to direct-current operation for 400-watt projection lamps at rated voltage (110 volts, 3140°K) burned base down.*

Type of Operation	$d\epsilon/dt$ L/W per Hour	$d\theta/dt$ Degrees K per Hour
A.C.	-0.058 ± 0.006	-0.33 ± 0.18
D.C.	$-0.059\pm .006$	$-0.69\pm .12$

current was given up and alternating-current burning used.

Later on, however, when the least-square analysis was applied to the data, it was found that the rather uncertain difference previously noted might be significant. Table V shows the result of the analysis carried out as previously described for the results presented in Table III. Note that the difference in decline of luminous efficiency is quite insignificant, but that apparently direct current produces a significantly more rapid decline in color temperature than alternating current. No satisfactory explanation of this result has been found. From the computed uncertainties it is to be expected to occur because of pure chance only about once in ten thousand such tests as were carried out.

D. Correlation between luminous efficiency and color temperature of individual lamps in the life test

It has been proposed that measurements of luminous efficiency might be made on lamp standards of color temperature as a check on their permanence. This proposal was based on two assumptions, first, that the precision with which luminous flux may be measured is considerably superior to that with which color temperature may be measured, and second, that there is a close correlation between the two quantities. The first assumption is not very well borne out by present practice because the superiority is slight. However, the second assumption has some justification, if referred to average values for groups of lamps as may be seen from Table III, though Table V illustrates a case of poor correlation.

Accordingly an attempt was made for a number of the individual lamps tested to correlate luminous efficiency with color temperature. The correlation found was quite poor, too poor to be of any use; there was little tendency for the lamps having rapidly decreasing luminous efficiency to have rapidly decreasing color temperature as well. This lack of correlation accords with the conclusion reached from theoretical considerations (see Section II) that both luminous efficiency and color temperature decline chiefly because of the brown deposit on the inside of the bulb, and that the former decrease corresponds to the average deposit all over the bulb, the latter to the deposit which forms on that part of the bulb in the line of sight. This view suggests that an improved correlation might be obtained by measuring apparent candle power in this one direction instead of mean spherical candle power, but the permanence of lamp standards of color temperature may now be so conveniently and reliably checked by color-temperature comparison with other similar standards set aside for that purpose that further work seems not to be justified.

E. Effects of seasoning the lamps

The changes which occur during seasoning are thought to be (1) expulsion of gases from the filament and a sintering of the filament accompanied by a decrease in resistance, and (2) a surface change of the filament causing a rise in color temperature. This surface change might be merely a decrease in area causing a rise in filament temperature, or it might be a change in surface character causing an increase of emissivity in the blue relative to that for red. The result with the projection lamps at 3100°K was unexpected because the current fell sharply indicating a sharp increase in filament resistance. The rise in color temperature was as expected. All lamps of the twenty seasoned at 3100°K agreed in these respects.

For the lamps constructed especially for photometric standards no rise in color temperature was found. The current, however, increased gradually throughout the 128-hour period of burning. This held true for each of the 10 lamps tested. This increase suggested that the lamps had not been thoroughly seasoned and that sintering of the filament was continuing throughout the entire period of burning at 2850°K. Details of construction and seasoning which might account for this result are not known.

F. Supplementary measurements of the brown deposit

The results of the tests thus far discussed suggest that the color-temperature impermanence of seasoned incandescent lamps is to be ascribed to brown deposit on the inside of the bulb. The amount of deposit in the line of sight even for the lamps burned for 700 hours is too small to be certainly detected by direct visual inspection although a dark deposit usually appears near the top of the lamps even before any certain decline in color temperature is noted. The question arises, therefore, whether there is really enough brown deposit in the line of sight to account for the decline in color temperature.

Accordingly the bulb of a lamp which had declined in color temperature about 20°K from 64 hours burning near 3140°K was removed and the spectral transmittance[13] of the deposit determined on the König-Martens spectrophotometer[14] both for the part of the bulb in the line of sight during the color-temperature measurements and for a part of the bulb near the top which seemed to bear the heaviest deposit. The spectrophotometer provided for unidirectional illumination and viewing, the direction of both being normal to the surface of the specimen.

Although the lamp was chosen because of the relative uniformity of the brown deposit on its bulb, this deposit was found to be sufficiently nonuniform to cause considerable trouble in making photometric settings. The results shown in Table VI are values adopted from a smooth curve drawn among the points representing the actual data. The last column in the table shows the seventh root of the transmittance of the heavy deposit. It is seen that these values agree very well with the transmittance of the deposit in the line of sight. This agreement indicates (1) that the nonuniformity of the deposit has not vitiated the results, (2) that the deposit at the top of the bulb is merely a thicker layer of the same material forming the deposit in the line of sight, and (3) that the deposit at the top of the bulb is seven times as thick as that on the bulb in the line of sight.

[13] For definition of this term, see J. O. S. A. **10**, 177 (1925).

[14] H. J. McNicholas, "Equipment for Routine Spectral Transmission and Reflection Measurements," Nat. Bur. Stand. J. Research **1**, 793 (1928); RP30.

TABLE VI. *Spectral transmittance of brown deposit on the bulb of a lamp whose color temperature had declined about 20°K during a 64-hour period of burning near 3140°K; illuminating beam and direction of view both nearly normal to surfaces of deposit.*

Wave-Length	Transmittance		Seventh Root of Transmittance at Top of Bulb
	In Line of Sight	At Top of Bulb	
440	0.848	0.306	0.843
480	.861	.354	.861
520	.874	.400	.876
560	.886	.437	.888
600	.897	.471	.897
640	.907	.502	.906
680	.916	.533	.914

By the red-blue-ratio method it is found that the deposit in the line of sight is sufficiently selective to reduce the color temperature of the lamp by 65°K near 3140, and that the heavy deposit is enough for a reduction of nearly 400°K. Since the observed decline in color temperature is only 20°K, it may be concluded that the selectivity of the brown deposit is ample to account for observed changes in color temperature.

In a sense, the measurements have proved too much; they indicate that three times the observed change ought to have been found. This might be caused by (1) a rise of 45°K in the color temperature of the filament from development of a severe "hot spot," or (2) to a possible division of the observed selectivity into about two-thirds from selective scattering and one-third from selective absorption. From Table II it is judged that the lamp would burn out from a "hot spot" severe enough to cause a rise of even half of 45°K; furthermore no "hot spots" were observed in the burn-out of these lamps; they burned normally with an occasional fusing of adjacent coils of the filament and continued loss of filament strength due to crystallization until an accidental jar broke the filament which would then crumble into many pieces. This rules out the first possibility.

Whether the deposit owes its measured selectivity to scattering or to absorption deserves further consideration. If the selectivity found by unidirectional illuminating and viewing normal to the surfaces of the deposit were entirely caused by selective scattering, such a deposit would be

expected to change the color temperature of the lamp not at all, since the short-wave-light which is scattered by the fragment of bulb in the spectrophotometer and lost by not entering the viewing slit is saved in the measurements of color temperature because the test plate is illuminated not only by light penetrating the bulb normally along the line of sight but also by light scattered from the deposit on other parts of the bulb. Accordingly an attempt was made to determine what part of the selectivity observed was caused by selective scattering and what part by selective absorption. What is required for measurement of the selectivity from selective absorption alone is spectral transmittance from unidirectional illumination counting in the measurement all light transmitted regardless of direction. The equivalent conditions, diffuse illumination and normal viewing, were tried by placing the sample over the sphere of the Priest-Lange reflectometer[15] and some evidence was obtained that a considerable part of the selectivity is ascribable to selective scattering but the nonuniformity of the sample introduced too much uncertainty in the measurement to permit any quantitative conclusion.

VII. Conclusions

The results of these tests may be summarized as follows:

1. Decrease of color temperature and luminous efficiency of seasoned incandescent lamps during burning at constant applied voltage is caused by the brown deposit forming on the inside of the bulb.

2. Color temperature of a lamp declines during such burning approximately in proportion to the twenty-third power of the filament temperature.

3. Luminous efficiency declines approximately in proportion to the thirty-first power of the filament temperature.

4. No good correlation between color temperature and luminous efficiency of individual lamps in the life test was found, nor is any to be expected since the two quantities depend differently on the distribution of the brown deposit over the inside surface of the bulb.

5. The spectral transmittance of the brown deposit has been found for two thicknesses, one deposit occurring at the side of a tubular bulb, the other, apparently about seven times as thick, occurring at the top of the bulb. The spectral selectivity of this deposit is probably to be ascribed partly to selective absorption and partly to selective scattering.

6. Lamps with tubular bulbs seem to have a more constant color temperature than those with pear-shaped bulbs probably because a greater proportion of the brown deposit on the tubular bulb occurs near the top out of the line of sight.

7. These results suggest that the permanence of lamp standards of color temperature would be increased several fold by building into the lamps a device for brushing off the deposit, but 400-watt projection lamps are already very permanent, decreasing in color temperature when burned at about 2850°K by an average of less than one-tenth of a degree per hour.

VIII. Acknowledgments

The operation of the lamps at constant voltage was carried out by R. S. Hunter, and the measurements of luminous flux were made by R. P. Teele and L. E. Barbrow. Dr. J. F. Meyer aided with his knowledge of incandescent lamps and his acquaintance with the extensive literature of the subject.

[15] I. G. Priest, "The Priest-Lange Reflectometer Applied to Nearly White Porcelain Enamels," Nat. Bur. Stand. J. Research **15**, 529 (1935); RP847.

Estimation of Chromaticity Differences and Nearest Color Temperature on the Standard 1931 ICI Colorimetric Coordinate System*

DEANE B. JUDD, *National Bureau of Standards, Washington, D. C.*

(Received September 23, 1936)

Estimation of chromaticity differences has been facilitated by the preparation of a standard mixture diagram showing by a group of ellipses the scales of perceptibility at the various parts of the diagram. The distances from the boundaries of the ellipses to their respective "centers" all correspond approximately to the same number (100) of "least perceptible differences." The estimation of nearest color temperature has been facilitated by the preparation of a mixture diagram on which is shown a family of straight lines intersecting the Planckian locus; each straight line corresponds approximately to the locus of points representing stimuli of chromaticity more closely resembling that of the Planckian radiator at the intersection than that of any other Planckian radiator.

CONTENTS

I. Introduction 421
II. Estimation of chromaticity differences 422
III. Nearest color temperature 423

I. INTRODUCTION

SINCE the adoption in 1931 by the International Commission on Illumination of a standard colorimetric coordinate system, by far the majority of fundamental colorimetric specifications have been expressed in that system.[1] A minor defect which is coming to be more keenly felt as the use of the system becomes more extended is the difficulty of interpreting differences in position on the standard mixture diagram in terms of degree and kind of chromaticity difference. The discovery of a projective transformation of the standard system whose Maxwell triangle yields nearly uniform chromaticity scales[2] is a step toward remedying this defect; on this triangle the length of any line is approximately proportional to the perceptibility of the chromaticity difference between the stimuli represented at the extremes of the line.

The substitution of the uniform chromaticity scale (UCS) system for the ICI system immediately suggests itself, but there are several reasons why this should not be done at present. In the first place, although the UCS system was derived from and is a valid representation of about nine-tenths of available data on sensibility to chromaticity differences, and although furthermore many divergent applications of it[3] have shown it to be a definite improvement in spacing over the ICI system, nevertheless the original data are neither sufficiently numerous nor sufficiently self-consistent to be considered definitely final. It is probable that further important improvements in spacing can be made. A second reason for adherence to the present standard ICI system is that the ICI system is more suited to the representation of photometric quantities; this may be seen from the fact that the ICI luminosity coefficients are 0, 1, 0, while the UCS luminosity coefficients are 0.08852, 0.11112, −0.09802. A third reason is that the applicability of data obtained by observing small apertures to observation of chromaticity differences between extended surface colors has not been demonstrated, nor have these data on small differences been shown to apply by simple addition to large chromaticity differences. The optimistic view that such applicability does not require demonstration has been fairly general, but extensive data which are being accumulated do not wholly justify this view. They indicate that the UCS triangle is applicable to small chromaticity differences between surface colors of large as well as small extent; furthermore, the application of the UCS triangle to large as well as small chromaticity differences between point

* Publication approved by the Director of the National Bureau of Standards of the U. S. Department of Commerce.

[1] D. B. Judd, "The 1931 I.C.I. Standard Observer and Coordinate System for Colorimetry," J. Opt. Soc. Am. **23**, 359 (1933).

[2] D. B. Judd, "A Maxwell Triangle Yielding Uniform Chromaticity Scales," Nat. Bur. Stand. J. Research **14**, 41 (1935) RP756; also J. Opt. Soc. Am. **25**, 24 (1935).

[3] See, for example, D. B. Judd, "A Method for Determining Whiteness of Paper," Paper Trade Journal, **100**, No. 21, TS40 (1935); also Tech. Assoc. Papers, Series XVIII, 392 (1935); and D. B. Judd, "A Method for Determining Whiteness of Paper, II," Tech. Assoc. Papers, Series XIX, 359 (1936); also Paper Trade Journal **103**, No. 8, TS38 (1936).

sources (signal lights) fails to result in any discrepancies not explainable by aberrations of the lens system of the eye; but there are data which suggest that the UCS triangle is not applicable to large chromaticity differences between extended surfaces. For these three reasons no departure from the standard 1931 ICI system is proposed at this time.

Accordingly, then, for research purposes many determinations in the colorimetry section of the National Bureau of Standards are being recorded in duplicate, once in the ICI system and once in the UCS system; but there are also many determinations which are expressed in the ICI system alone. For such determinations some of the advantages of the UCS system may be obtained if there is available a graphical representation of the relative scales of the two systems. The present paper serves to give two such representations; one of them shows how the scale of perceptibility varies over the ICI mixture diagram on the assumption that the UCS system gives accurately uniform chromaticity scales as it was intended to do, the other shows, on the same assumption, iso-temperature lines to aid in the estimation of nearest color temperature for near-Planckian colors.[4]

II. Estimation of Chromaticity Differences

The equilateral Maxwell triangle of the ICI system and all three of the mixture diagrams obtained by plotting in rectangular coordinates one of the ICI trilinear coordinates (x, y, z) against another fail to array color stimuli in such a way that equal chromaticity differences are indicated by lines of equal length. It is therefore impossible to give any single reliable perceptibility scale of chromaticity which will apply to the whole diagram as a scale of miles applies to a map. Instead of a single perceptibility scale there is required for a given direction on the mixture diagram a different scale for each different portion of the diagram, and for each point a different scale is, in general, required for each direction. In an effort to provide such a complete set of approximate perceptibility scales for the ICI mixture diagram a group of thirty tangent circles on the UCS system have been transferred to the (x, y) plot of the ICI system where they appear as tangent ellipses of various sizes, eccentricities and orientations; see Fig. 1. The distances from points on the boundaries of these ellipses to the point within them, representing the transformed centers of the original circles, give the desired scales of perceptibility; one one-hundredth of each such distance refers to a chromaticity difference just perceptible with certainty under moderately good experimental conditions (careful monocular observation of a 6° circular field divided along a diameter yielding retinal illuminations between 200 and 1000 photons).

The indicated maximum variation in the scales of perceptibility is about 14 to 1; the distance from point to boundary along the major axis of the largest ellipse ($x=0.23$, $y=0.77$) is about 0.25, while the smallest such distance parallel to a minor axis (see ellipse inclosing $x=0.43$, $y=0.12$) is about 0.018. In most practical uses of the ICI system, however, the maximum variation is only about 5 to 1 because green samples whose colors are so saturated as to fall within the largest ellipses are rare.

The following example will show how Fig. 1 may be used as an aid in the estimation of chromaticity differences. Suppose it be required to know how much the dominant wave-length of a yellow signal light could be changed from 580 mμ and still keep the chromaticity difference no more perceptible than that between the signal at 580 mμ and a bare incandescent lamp at a color temperature of 2360°K. Since signals produced by yellow glasses may be closely matched by one part or another of the spectrum, itself, the spectrum may be used for this estimation. The scale in the direction of 580 mμ from 2360°K is about 0.13 per hundred "least perceptible differences" (LPDs); the scale along the spectrum near 580 mμ is about 0.04 per 100 LPDs. The distance on the diagram along the spectrum locus equivalent to the distance between 2360°K and 580 mμ is: $0.075(0.04/0.13) = 0.023$. This corresponds to slightly less than 3 mμ near 580, hence it would be said that the diagram indicates that both the difference between 2360°K and 580 mμ and that between

[4] For these terms and concepts, see R. Davis, "A Correlated Color Temperature for Illuminants," Nat. Bur. Stand. J. Research **7**, 659 (1931); RP365.

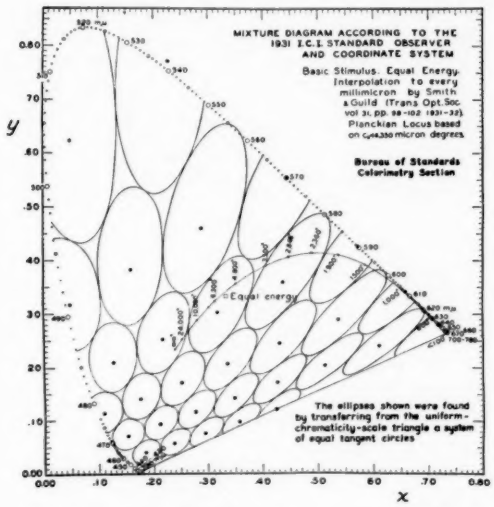


FIG. 1. Perceptibility scales for the standard 1931 ICI colorimetric coordinate system. The distances from points on the boundary of each ellipse to the indicated point within it all correspond approximately to one hundred times the chromaticity difference just perceptible with certainty under moderately good observing conditions.

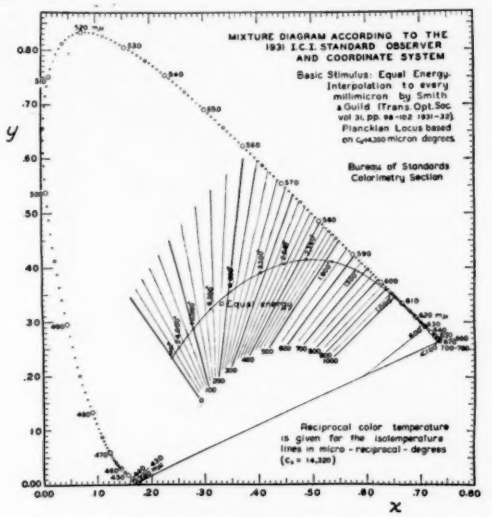


FIG. 2. Iso-temperature lines shown on the standard 1931 ICI colorimetric coordinate system. Color stimuli specified on an iso-temperature line give closer chromaticity matches with the indicated Planckian color than with Planckian colors at neighboring color temperatures.

580 and 583 mμ are about 60 times that least perceptible under good experimental conditions.

It is worth noting that a yellow signal of 580 mμ would resemble a bare vacuum-tungsten lamp sufficiently to be frequently confusible; this accounts for the non-use of such a signal. For example, railway signal yellow is a reddish yellow of dominant wave-length greater than 588 mμ.[5]

This example also serves to emphasize that the size of the LPD varies widely with experimental conditions. The unfavorable conditions often arising in the observation of railway signals (momentary viewing of point source at a retinal illumination of less than 100 photons with no comparison light) require perhaps fifty to one hundred times as large chromaticity differences for discrimination as are required under good laboratory conditions.

III. Nearest Color Temperature

An important practical application of the UCS system is its use in finding from any series of colors the one most resembling any neighboring color of the same brilliance. In this system the method, of course, is to draw the shortest line from the locus of the series to the point representing the neighboring color. This is the geometric equivalent of the method of observation on which many plans of color grading are based. The characteristic of these plans is that they pay attention to chromaticity differences along a series of color standards, and in the usual case that no member of the series gives a perfect chromaticity match, the nearest match is set and the residual difference neglected. In this way a problem which, strictly, is two-dimensional finds a one-dimensional solution that is close enough to be of practical use.

In finding nearest color from a series the UCS system may be used directly. The locus of the series is plotted together with the point representing the neighboring chromaticity, and the shortest line through the point to the locus, usually the normal, gives the nearest chromaticity of the series. It is also possible to carry out the operation once for all by drawing on the UCS triangle a number of normals to the locus of the series completely covering at sufficiently small intervals the areas of the triangle near the

[5] Association of American Railroads, Signal Section Specification 69–35, Signal glasses (exclusive of hand lantern globes). *Manual of Signal Section*, AAR, Part 136 (1935).

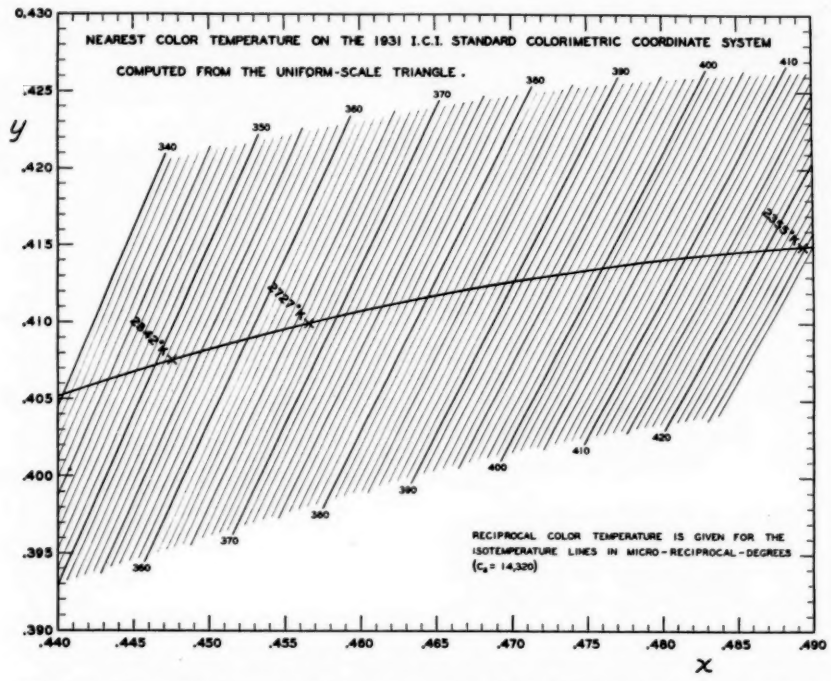

NEAREST COLOR TEMPERATURE ON THE 1931 I.C.I. STANDARD COLORIMETRIC COORDINATE SYSTEM
COMPUTED FROM THE UNIFORM-SCALE TRIANGLE.
RECIPROCAL COLOR TEMPERATURE IS GIVEN FOR THE ISOTEMPERATURE LINES IN MICRO-RECIPROCAL-DEGREES (C_2 = 14,320)
2942°K
2727°K
2539°K
y
x

FIG. 3. Iso-temperature lines shown for a small section of Fig. 2. This region includes the most-used portion of the color-temperature scale for incandescent lamps.

locus. These normals can then be transformed to any other desired system and the determination of nearest chromaticity carried on entirely outside the UCS system. This has been done by McNicholas in the estimation of nearest chromaticity of vegetable oils on both the wavelength scale and on the scale composed of combinations of Lovibond 35-yellow with various Lovibond red glasses.[6] Fig. 2 shows a similar series of normals to the Planckian locus transferred to the standard 1931 ICI system.[7] By means of these transferred normals (iso-temperature lines) nearest color temperature estimated according to the UCS system may be approximately read from specifications in the ICI system. Fig. 3 shows a small section of Fig. 2 to an enlarged scale with one normal drawn in for each microreciprocal degree (μrd); as pointed out by Priest[8] who proposed the use of reciprocal color temperature, "a difference of one microreciprocal degree is fairly representative of the doubtfully perceptible difference in chromaticity under the most favorable conditions of observation." The large drawing from which this reproduction was made has been used in connection with a check of the present color-temperature scale[9] with blue filters whose use resulted in near-Planckian colors. The original drawing was found to reproduce estimates of nearest color

[6] H. J. McNicholas, "Color and Spectral Transmittance of Vegetable Oils," Nat. Bur. Stand. J. Research **15**, 99 (1935); RP 815; also Oil & Soap **12**, 167 (1935); see his Fig. 8.

[7] Fig. 2 is a reproduction of a figure previously used (see reference 1). Since that time a new color temperature scale has been adopted (see reference 9) and coincidentally the value of C_2 used was changed from 14,350 to 14,320 micron degrees to conform with the International Temperature Scale. This change in C_2 makes only 0.2 percent change in the temperature (or reciprocal temperature) and is too small to be shown in Fig. 2. For this reason the old diagram has been used. In Fig. 3, however, the change (about 0.8 μrd) could be easily seen.

In Table II, the values of color temperature listed were taken from Davis' paper in which the older value of 14,350 was used for C_2.

[8] I. G. Priest, "A Proposed Scale for Use in Specifying the Chromaticity of Incandescent Illuminants and Various Phases of Daylight," J. Opt. Soc. Am. **23**, 41 (1933).

[9] H. T. Wensel, D. B. Judd and Wm. F. Roester, "Establishment of a Scale of Color Temperature," Nat. Bur. Stand. J. Research **12**, 527 (1934); RP677.

temperature found directly from the UCS system with a maximum error less than 0.1 μrd. It will be noted that the normals extend only over the region immediately in the neighborhood of the Planckian locus. The experimental setting of nearest color temperature when a considerable chromaticity difference exists between the illuminant in question and the Planckian radiator at any temperature becomes increasingly difficult and ambiguous as the chromaticity difference is increased. It is doubtful whether the solution given by the UCS system is valid much further than the normals have been extended on Fig. 3; and Fig. 2, which shows them extended considerably further, may be criticized because it seems to make the false implication that a definite, unambiguous nearest color temperature can be found experimentally for illuminants having colors very notably non-Planckian.

To facilitate the construction of large-scale plots for accurate work such as Fig. 3 there have been included in Table I the trilinear coordinates of the Planckian locus on the ICI system for the reciprocal color temperatures represented by the transformed normals on Fig. 2, and the reciprocals of the slopes of these transformed normals. The reciprocals of the slopes are given instead of the slopes, themselves, for convenience because, as may be seen from Fig. 2, the slope becomes infinite in the neighborhood of 160 μrd, but is nowhere zero. Interpolation between the values given in Table I will yield points on the Planckian locus and also the slopes of the corresponding iso-temperature lines.

Table II shows a comparison of conjunctive wave-length, λj, referring to correlated color temperature computed by Davis' empirical method with that referring to nearest color temperature from the UCS system. The former

TABLE I. *Trilinear coordinates of Planckian stimuli and reciprocal of slopes of the iso-temperature lines as a function of reciprocal color temperature.*

$1/\theta$ $C_2=14{,}320$ μrd	Trilinear coordinates of Planckian locus		Reciprocal of the slopes of the iso-temperature lines
	x	y	
0	0.2399	0.2342	−0.672
20	.2458	.2429	− .611
40	.2526	.2524	− .536
60	.2605	.2629	− .450
80	.2693	.2743	− .359
100	.2794	.2869	− .266
120	.2909	.3003	− .175
140	.3037	.3141	− .089
160	.3171	.3272	− .012
180	.3306	.3396	+ .057
200	.3445	.3512	+ .120
220	.3583	.3619	+ .175
240	.3723	.3717	+ .227
260	.3865	.3805	+ .275
280	.4006	.3883	+ .319
300	.4142	.3950	+ .361
320	.4273	.4006	+ .400
340	.4403	.4052	+ .437
360	.4526	.4089	+ .472
380	.4646	.4117	+ .505
400	.4760	.4136	+ .535
420	.4868	.4148	+ .563
440	.4973	.4153	+ .590
460	.5072	.4152	+ .615
480	.5167	.4146	+ .639
500	.5256	.4136	+ .661
550	.5460	.4093	+ .713
600	.5638	.4033	+ .762
650	.5797	.3962	+ .808
700	.5941	.3886	+ .849
800	.6183	.3731	+ .923
900	.6370	.3587	+ .983
1000	.6521	.3451	+1.036

TABLE II. *Comparison of conjunctive wave-length referring to nearest color temperature via the UCS system with that referring to Davis' correlated color temperature.*

θ $C_2=14{,}350$ deg K	$1/\theta$ $C_2=14{,}320$ μrd	Conjunctive wave-length, λj, in mμ	
		Nearest color temperature via UCS system	Davis correlated color temperature
1,803	555.8	588.4	586.5
2,011	498.3	585.6	584.2
2,532	395.8	580.2	579.5
3,053	328.2	575.4	575.5
3,585	279.5	570.9	572.9
4,116	243.5	567.2	570.2
4,663	214.9	563.0	568.8
5,222	191.9	559.1	567.0
5,781	173.3	555.5	565.8
6,359	157.6	549.8	564.8
6,945	144.3	546.8	564.0
7,549	132.7	542.0	563.2
8,181	122.5	537.7	562.6
8,826	113.5	533.5	562.1
9,483	105.7	530.0	561.4
10,187	98.4	526.2	561.0
10,876	92.1	524.5	561.2
11,587	86.5	521.7	561.0
13,161	76.1	517.0	560.3
14,817	67.6	514.4	560.3

were taken from Davis' Table VI,[4] the latter were found by extending the iso-temperature lines on the original large-scale plot of Fig. 2 until they cut the spectrum locus. It may be noted that the agreement is good down to 250 *μrd* (up to 4000°K) but is rather poor at lower reciprocal color temperatures where the validity of the UCS system is best established.[2]